JHEP

Published for SISSA by Springer

Received: January 25, 2019
Revised: April 1, 2019
Accepted: April 1, 2019
Published: April 15, 2019

A Monte Carlo global analysis of the Standard Model Effective Field Theory: the top quark sector

Nathan P. Hartland,[a,b] **Fabio Maltoni,**[c,d] **Emanuele R. Nocera,**[b,e] **Juan Rojo,**[a,b] **Emma Slade,**[f] **Eleni Vryonidou**[g] **and Cen Zhang**[h]

[a] *Department of Physics and Astronomy, Vrije Universiteit Amsterdam, NL-1081 HV Amsterdam, The Netherlands*

[b] *Nikhef Theory Group, Science Park 105, 1098 XG Amsterdam, The Netherlands*

[c] *Centre for Cosmology, Particle Physics and Phenomenology (CP3), Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium*

[d] *Dipartimento di Fisica e Astronomia, Università di Bologna and INFN, Sezione di Bologna, via Irnerio 46, 40126 Bologna, Italy*

[e] *The Higgs Centre for Theoretical Physics, University of Edinburgh, JCMB, KB, Mayfield Rd, Edinburgh EH9 3FD, Scotland*

[f] *Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, United Kingdom*

[g] *Theoretical Physics Department, CERN, CH-1211 Geneva, Switzerland*

[h] *Institute of High Energy Physics, and School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*

E-mail: nphartland@gmail.com, Fabio.Maltoni@uclouvain.be, emanuele.nocera@physics.ox.ac.uk, j.rojo@vu.nl, emma.slade@physics.ox.ac.uk, evryonidou@gmail.com, cenzhang@ihep.ac.cn

Abstract: We present a novel framework for carrying out global analyses of the Standard Model Effective Field Theory (SMEFT) at dimension-six: SMEFiT. This approach is based on the Monte Carlo replica method for deriving a faithful estimate of the experimental and theoretical uncertainties and enables one to construct the probability distribution in the space of the SMEFT degrees of freedom. As a proof of concept of the SMEFiT methodology, we present a first study of the constraints on the SMEFT provided by top quark production measurements from the LHC. Our analysis includes more than 30 independent measurements from 10 different processes at $\sqrt{s} = 8$ and 13 TeV such as inclusive $t\bar{t}$ and single-top production and the associated production of top quarks with weak vector bosons and the Higgs boson. State-of-the-art theoretical calculations are adopted both for the Standard Model and for the SMEFT contributions, where in the latter case NLO

https://doi.org/10.1007/JHEP04(2019)100

JHEP04(2019)100

QCD corrections are included for the majority of processes. We derive bounds for the 34 degrees of freedom relevant for the interpretation of the LHC top quark data and compare these bounds with previously reported constraints. Our study illustrates the significant potential of LHC precision measurements to constrain physics beyond the Standard Model in a model-independent way, and paves the way towards a global analysis of the SMEFT.

Keywords: Beyond Standard Model, Effective Field Theories, Perturbative QCD

ArXiv ePrint: 1901.05965

JHEP04(2019)100

Contents

1 Introduction

The Large Hadron Collider (LHC) is pursuing an extensive program of direct searches for physics beyond the Standard Model (BSM) by exploiting its unique reach in energy. Whilst these searches have not yet returned any convincing evidence for BSM physics, only a small fraction of the final LHC dataset has been analysed so far, and ample room for surprises remains. A complementary approach to the searches for direct production of new particles is

JHEP04(2019)100

that of *indirect* BSM searches, where precise measurements of total cross-sections and differential distributions are compared to Standard Model (SM) predictions with the hope to uncover glimpses of BSM dynamics in the interactions between SM particles. For instance, if new particles are too heavy to be directly produced at the LHC, they could still leave imprints in the kinematical distributions of the SM particles via interference or virtual effects.

A powerful framework to identify, constrain, and parametrise potential deviations with respect to the SM predictions in a model-independent way is the Standard Model Effective Field Theory (SMEFT) [1–3]. In this framework, the effects of BSM dynamics at high scales $E \simeq \Lambda$ are parametrised for $E \ll \Lambda$ in terms of higher-dimensional (irrelevant) operators built up from the SM fields and respecting symmetries such as gauge and Lorentz symmetry. This approach is robust and general, since one can construct non-redundant bases of independent operators at any given mass dimension ($\hbar = c = 1$) that can then be systematically matched to explicit ultraviolet-complete scenarios for their interpretation at any order in $1/\Lambda$.

Analysing experimental data in the SMEFT framework is non trivial; even restricting oneself to operators that conserve baryon and lepton number [3], one ends up with $N_{\rm op} = 59$ operators at dimension six for one generation, growing to more than 2000 in absence of flavour assumptions. This implies that global and model-independent SMEFT analyses need to explore a complicated parameter space with a large number of degenerate ("flat") directions and local minima.

In this context, the wealth of precision measurements presented by the LHC collaborations in recent years, together with the significant progress in the corresponding theoretical calculations and modelling of collider processes, has motivated many groups to pursue (partial) SMEFT analyses of the LHC data [4–21] complemented often with input from lower-energy experiments such as the LEP electroweak precision tests. In these fits, constraints on the SMEFT operators can be provided not only by "traditional" processes such as electroweak gauge boson and Higgs production, but also by other high-p_T processes such as jet and top quark production. Interestingly, even when only considering electroweak processes, these constraints are comparable or even superior to those provided by LEP [4, 22]. Indeed, SMEFT corrections often grow quadratically with the energy and thus directly benefit from the large kinematic reach, up to several TeV, provided by present and future LHC measurements.

From the methodological point of view, a global fit of the SMEFT from LHC measurements requires combining state-of-the-art theoretical calculations (in the SM and in the SMEFT) with a wide variety of experimental cross-sections and distributions. This should be accomplished by means of a robust statistical analysis allowing for the reliable estimation of all sources of uncertainty and for the minimisation of procedural and theoretical biases. SMEFT fits therefore represent, conceptually, a similar problem to that arising in the global QCD analysis of the quark and gluon structure of the proton in terms of parton distribution functions (PDFs) [23–25]. By exploiting these conceptual similarities, in this work we develop a novel strategy for global SMEFT analyses inspired by the NNPDF framework, successfully applied to the determination of the parton distributions of the proton [26–36] and of hadron fragmentation functions [37, 38]. This approach, which

JHEP04(2019)100

we denote by SMEFiT, combines the generation of Monte Carlo (MC) replicas, to estimate and propagate uncertainties, with cross-validation to prevent over-fitting.

As a proof of concept of the SMEFiT methodology, we apply it here for the first time to the detailed study of top quark production at the LHC in the SMEFT framework at dimension six. The top quark, the only fermion with an $\mathcal{O}(1)$ Yukawa coupling, plays a privileged role in most BSM scenarios aiming to explain the origin of electroweak symmetry breaking and stabilise the weak scale. From the experimental data point of view, a global SMEFT analysis of top quark production at the LHC is motivated by the large number of precision measurements at $\sqrt{s} = 7, 8$ and 13 TeV that have become available recently. This data includes total rates and differential distributions in inclusive $t\bar{t}$ and single-top production, associated production of top quarks with vector bosons and the Higgs boson, and helicity fractions in top quark decay. The wealth of data collected by the LHC is mirrored by the advancements on the theoretical side, where significant progress in higher order calculations in the top quark sector has been achieved. This is true both from the SM point of view, with the calculation of NNLO QCD and NLO electroweak corrections for inclusive top quark pair and single top production, as well as from the SMEFT side. In the latter case, LO calculations are now automatised in codes such as `MadGraph5_aMC@NLO` [39] within a framework agreed within the LHC Top WG [10], and NLO QCD corrections have been presented for a continuously growing number of processes.

Several SMEFT analyses of the top quark sector have been presented based on either hadron collider [40–45] or lepton collider [46] processes, in the latter case also considering the sensitivity of future machines such as the International Linear Collider (ILC). The top quark sector of the SMEFT has been in particular studied by the `TopFitter` collaboration [15, 47, 48]. Our analysis exhibits several improvements as compared to the available studies, allowing us to assess the impact of several important aspects in the fit. First, we include a broader range of input experimental measurements from different processes, which allow us to constrain a larger number of SMEFT operators. Second, we include the NLO QCD corrections to the SMEFT contributions. This entails an improved accuracy and a reduction of the theory systematic errors. Third, we always compute both the leading linear ($\mathcal{O}(\Lambda^{-2})$) and the subleading quadratic ($\mathcal{O}(\Lambda^{-4})$) contributions to the SMEFT predictions, so that effects of including or not the quadratic terms can be systematically studied. Fourth, our methodology avoids any assumption about the specific profile of the χ^2 function and in particular we do not rely on any quadratic approximation for error propagation.

By exploiting the SMEFiT methodology, here we derive the probability distribution in the space of SMEFT Wilson coefficients that follows from all available top quark production cross-sections. We study the impact of individual processes on the SMEFT parameter space and the role of higher order corrections, such as NLO QCD and the SMEFT $\mathcal{O}(\Lambda^{-4})$ corrections. In general, we find that higher order effects are non-negligible and can significantly affect the results. We also quantify the correlations between the operators, and compare the bounds derived here with previous constraints reported in the literature. Our analysis illustrates the significant potential of LHC precision measurements to constrain, and possibly identify, BSM physics in a model-independent way.

JHEP04(2019)100

The outline of this paper is the following. In section 2 we summarise the SMEFT description of the top quark sector at dimension six and introduce our choice of operator basis for the fit. In section 3 we describe the experimental measurements of top quark production at the LHC which are used to constrain the SMEFT operators and the settings of the corresponding theoretical calculations of the SM and SMEFT cross-sections. The SMEFiT methodology is presented in section 4, where it is validated by means of closure tests. The main results of this work are presented in section 5, where we determine the confidence level intervals for the coefficients of the $N_{\rm op} = 34$ SMEFT operators and their correlations, and compare them with the bounds reported in previous studies. In section 6 we summarise our main conclusions and outline possible directions for generalising our analysis to other processes.

2 The SMEFT in the top quark sector

In this section we describe the theoretical formalism that will be adopted in this work to interpret the LHC top quark production data within the SMEFT framework. First, we provide an introduction to the SMEFT, focusing on those operators that affect the description of the top quark sector. Then, we define the degrees of freedom that are more relevant to studying top quark production at the LHC. Operators that do not involve top quarks and their constraints are also briefly discussed. We finally describe our theory calculations at NLO QCD accuracy, and comment on some additional aspects of the SMEFT formalism relevant for this study.

2.1 The SMEFT framework

Let us begin by reviewing the SMEFT formalism [2, 49], with emphasis on its description of the top quark sector. As mentioned in the introduction, the effects of new heavy BSM particles with typical mass scale $M \simeq \Lambda$ can under general conditions be parametrised at lower energies $E \ll \Lambda$ in a model-independent way in terms of a basis of higher-dimensional operators constructed from the SM fields and their symmetries. The resulting effective Lagrangian then admits the following power expansion

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_{i}^{N_{d6}} \frac{c_i}{\Lambda^2} \mathcal{O}_i^{(6)} + \sum_{j}^{N_{d8}} \frac{b_j}{\Lambda^4} \mathcal{O}_j^{(8)} + \dots \,, \tag{2.1}$$

where $\mathcal{L}_{\rm SM}$ is the SM Lagrangian, and $\{\mathcal{O}_i^{(6)}\}$ and $\{\mathcal{O}_j^{(8)}\}$ stand for the elements of the operator basis of mass-dimension $d = 6$ and $d = 8$, respectively. Operators with $d = 5$ and $d = 7$, which violate lepton and/or baryon number conservation [50, 51], are not considered here. Whilst the choice of operator basis used in eq. (2.1) is not unique, it is possible to relate the results obtained in different bases [52]. In this work we adopt the Warsaw basis for $\{\mathcal{O}_i^{(6)}\}$ [3], and neglect effects arising from operators with mass dimension $d \geq 8$.

For specific UV completions, the Wilson coefficients $\{c_i\}$ in eq. (2.1) can be evaluated in terms of the parameters of the BSM theory, such as its coupling constants and masses. However, in a bottom-up approach, they are *a priori* free parameters and they need to

JHEP04(2019)100

be constrained from experimental data. In general, the effects of the dimension-6 SMEFT operators in a given observable, such as cross-sections at the LHC, differential distributions, or other pseudo-observables, can be written as follows:

$$\sigma = \sigma_{\rm SM} + \sum_{i}^{N_{d6}} \kappa_i \frac{c_i}{\Lambda^2} + \sum_{i,j}^{N_{d6}} \widetilde{\kappa}_{ij} \frac{c_i c_j}{\Lambda^4} \,, \tag{2.2}$$

where $\sigma_{\rm SM}$ indicates the SM prediction and the Wilson coefficients c_i are considered to be real for simplicity.

In eq. (2.2), the second term arises from operators interfering with the SM amplitude. The resulting $\mathcal{O}\left(\Lambda^{-2}\right)$ corrections to the SM cross-sections represent formally the dominant correction, though in many cases they can be subleading for different reasons. The third term in eq. (2.2), representing $\mathcal{O}\left(\Lambda^{-4}\right)$ effects, arises from the squared amplitudes of the SMEFT operators, irrespectively of whether or not the dimension-6 operators interfere with the SM diagrams. In principle, this second term may not need to be included, depending on if the truncation at $\mathcal{O}\left(\Lambda^{-2}\right)$ order is done at the Lagrangian or the cross section level, but in practice there are often valid reasons to include them in the calculation. We will discuss in more details the impact of these $\mathcal{O}\left(\Lambda^{-4}\right)$ corrections at the end of this section.

An important aspect of any SMEFT analysis is the need to include all relevant operators that contribute to the processes whose data is used as input to the fit. Only in this way can the SMEFT retain its model and basis independence. However, unless specific scenarios are adopted, the number of non-redundant operators N_{d6} becomes unfeasibly large: 59 for one generation of fermions [3] and 2499 for three [53]. This implies that a global SMEFT fit, even if restricted to dimension-6 operators, will have to explore a huge parameter space with potentially a large number of flat (degenerate) directions.

Due to the above consideration, in this work we follow closely the strategy documented in the LHC Top Quark Working Group note [10]. In particular, we adopt the Minimal Flavour Violation (MFV) hypothesis [54] in the quark sector as the baseline scenario. We further assume that the Cabibbo-Kobayashi-Maskawa (CKM) matrix is diagonal, and that the Yukawa couplings are nonzero only for the top and bottom quarks. In other words, we impose a $\mathrm{U}(2)_q \times \mathrm{U}(2)_u \times \mathrm{U}(2)_d$ flavour symmetry among the first two generations. In addition, we restrict ourselves to the CP-even operators only, and focus on those operators that induce modifications in the interactions of the top quark with other SM fields. As we will now show, under the above assumptions, we will explore the parameter space associated to the $N_{\rm op} = 34$ linear combinations of dimension-6 operators that are relevant for the description of the top quark sector. Following ref. [10], we will then define the specific degrees of freedom relevant for the interpretation of top quark measurements.

2.2 The top quark sector of the SMEFT

Given the scope of this study, we will consider here only those dimension-6 operators that affect the production and decay of top quarks at the LHC through the modifications of their couplings to other SM fields. Following ref. [10], we adopt the Warsaw basis [3] of non-redundant, gauge-invariant dimension-six operators, and then we define the specific

JHEP04(2019)100

degrees of freedom relevant for each measurement. These degrees of freedom are linear combinations of the Warsaw-basis operator coefficients, which appear in the interference with SM amplitudes, and in interactions with physical fields after electroweak symmetry breaking. These combinations are then aligned with physically relevant directions of the SMEFT parameter space. They represent the maximal information that can be extracted from measuring a certain process. The rationale for using them in a global fit instead of the basis operator coefficients directly is that they may reduce the number of relevant parameters and unconstrained combinations.

Since we only consider here those operators which contain at least one top quark under the assumed flavour symmetries, we are implicitly assuming that other operators affecting the considered processes are well constrained from measurements of other processes that do not involve top quarks. This assumption may not always be justified, but it is helpful for a better understanding of the top quark sector, and also for setting up the scope of this work. Without this assumption, it is likely that one would have to resort to a much more global analysis, including all currently available data, which goes beyond the scope of the present analysis. We will discuss explicitly how in our case this assumption is justified in the next subsection.

We are now ready to define the relevant degrees of freedom that will be used in this analysis in terms of the dimension-6 operators of the Warsaw basis. The complete set of degrees of freedom can be found in ref. [10], and for completeness we collect in appendix A the definitions and conventions that will be adopted in the following. To begin with, concerning the operators involving four heavy quarks (that is, either a right-handed top t, or a right-handed bottom b, or a left-handed top-bottom doublet Q), we define the following degrees of freedom:

$$\begin{aligned} c^1_{QQ} &\equiv 2C^{1(3333)}_{qq} - \frac{2}{3}C^{3(3333)}_{qq}, \\ c^8_{QQ} &\equiv 8C^{3(3333)}_{qq}, \end{aligned} \qquad \begin{aligned} c^1_{Qt} &\equiv C^{1(3333)}_{qu}, \\ c^8_{Qt} &\equiv C^{8(3333)}_{qu}, \\ c^1_{Qb} &\equiv C^{1(3333)}_{qd}, \\ c^8_{Qb} &\equiv C^{8(3333)}_{qd}, \end{aligned} \qquad c^1_{tt} \equiv C^{1(3333)}_{uu}, \qquad \begin{aligned} c^1_{tb} &\equiv C^{1(3333)}_{ud}, \\ c^8_{tb} &\equiv C^{8(3333)}_{ud}, \end{aligned} \tag{2.3}$$

and in addition we also have

$$c^1_{QtQb} \equiv \mathrm{Re}\{C^{1(3333)}_{quqd}\}, \qquad c^8_{QtQb} \equiv \mathrm{Re}\{C^{8(3333)}_{quqd}\}\,, \tag{2.4}$$

in terms of the Warsaw-basis operators listed in eq. (A.1). We note that the imaginary parts of the last two operators are CP-odd, and therefore are not included here since we restrict ourselves to CP-conserving operators.

Note that in eq. (2.3) all four flavour indices in these degrees of freedom correspond to the third (heavy) quark generation. For example, the degree of freedom labelled as c^1_{QQ} in eq. (2.3) is constructed from a linear combination of the $\mathcal{O}^{1(ijkl)}_{qq}$ and $\mathcal{O}^{3(ijkl)}_{qq}$ operators for which $i = j = k = l = 3$. From eq. (2.3) we see that within the specific flavour assumptions adopted here there are 11 operators involving four heavy quarks. These operators can only

JHEP04(2019)100

be constrained from processes involving four heavy quarks in the final state, such as four-top quark production or $t\bar{t}b\bar{b}$ production, as we will discuss below in section 3.4.

Concerning the dimension-6 operators of the Warsaw basis involving two light quarks and two heavy quarks, see the list in eq. (A.2), we first note that operators involving a light-quark scalar or tensor current are vetoed by the flavour assumptions adopted here. On the other hand, vector-like interactions such as $\bar{L}L\bar{L}L$, $\bar{L}L\bar{R}R$, and $\bar{R}R\bar{R}R$ type operators are allowed by our flavour scenario. We can therefore define the following degrees of freedom in terms of two-light-two-heavy operators:

$$\begin{aligned}
c_{Qq}^{1,1} &\equiv C_{qq}^{1(ii33)} + \frac{1}{6}C_{qq}^{1(i33i)} + \frac{1}{2}C_{qq}^{3(i33i)}, & c_{tu}^{1} &\equiv C_{uu}^{(ii33)} + \frac{1}{3}C_{uu}^{(i33i)}, & c_{tq}^{1} &\equiv C_{qu}^{1(ii33)}, \\
c_{Qq}^{3,1} &\equiv C_{qq}^{3(ii33)} + \frac{1}{6}(C_{qq}^{1(i33i)} - C_{qq}^{3(i33i)}), & c_{tu}^{8} &\equiv 2C_{uu}^{(i33i)}, & c_{Qu}^{1} &\equiv C_{qu}^{1(33ii)}, \\
c_{Qq}^{1,8} &\equiv C_{qq}^{1(i33i)} + 3C_{qq}^{3(i33i)}, & c_{td}^{1} &\equiv C_{ud}^{1(33ii)}, & c_{Qd}^{1} &\equiv C_{qd}^{1(33ii)}, \\
c_{Qq}^{3,8} &\equiv C_{qq}^{1(i33i)} - C_{qq}^{3(i33i)}, & c_{td}^{8} &\equiv C_{ud}^{8(33ii)}, & c_{tq}^{8} &\equiv C_{qu}^{8(ii33)}, \\
& & & & c_{Qu}^{8} &\equiv C_{qu}^{8(33ii)}, \\
& & & & c_{Qd}^{8} &\equiv C_{qd}^{8(33ii)},
\end{aligned} \tag{2.5}$$

where i corresponds to a light quark index, that is, it is either 1 or 2, and recall that the first two generations are massless and thus exhibit an SU(2) flavour symmetry. For these degrees of freedom involving two heavy quarks and two light quarks, we therefore end up with 14 independent coefficients. These degrees of freedom can be constrained by processes such as inclusive $t\bar{t}$, through the quark-antiquark component of the initial state, as well as by $t\bar{t}$ production in association with gauge vector bosons. The SU(2) triplet degrees of freedom can also be constrained by single top processes.

Finally, we need to take into account the degrees of freedom involving operators built from two heavy quarks and bosonic fields, including the Higgs field, namely those listed in eq. (A.2). For these operators, the following combinations are defined:

$$\begin{aligned}
c_{t\varphi} &\equiv \mathrm{Re}\{C_{u\varphi}^{(33)}\}, & c_{\varphi Q}^{-} &\equiv C_{\varphi q}^{1(33)} - C_{\varphi q}^{3(33)}, & c_{tW} &\equiv \mathrm{Re}\{C_{uW}^{(33)}\}, \\
& & c_{\varphi Q}^{3} &\equiv C_{\varphi q}^{3(33)}, & c_{tZ} &\equiv \mathrm{Re}\{-s_W C_{uB}^{(33)} + c_W C_{uW}^{(33)}\}, \\
& & c_{\varphi t} &\equiv C_{\varphi u}^{(33)}, & c_{bW} &\equiv \mathrm{Re}\{C_{dW}^{(33)}\}, \\
& & c_{\varphi tb} &\equiv \mathrm{Re}\{C_{\varphi ud}^{(33)}\}, & c_{tG} &\equiv \mathrm{Re}\{C_{uG}^{(33)}\}.
\end{aligned} \tag{2.6}$$

We see for example that the c_{tZ} degree of freedom is a combination of the $\mathcal{O}_{uB}^{ij}$ and $\mathcal{O}_{uW}^{ij}$ operators with $i = j = 3$, weighted by the sine and the cosine of the Weinberg angle respectively. Since we account here only for CP-conserving effects, the imaginary parts of the last five coefficients, being CP-odd, will not be included.

Note that there are two additional degrees of freedom that fall into the same category (two heavy quark fields plus bosonic fields), but they are not independent from those defined above. First of all, we have the combination of $O_{\varphi q}^{1,3(33)}$ operators that modifies the SM coupling of the b quark to the Z boson, defined as

$$c_{\varphi Q}^{+} \equiv C_{\varphi q}^{3(33)} + C_{\varphi q}^{1(33)} = c_{\varphi Q}^{-} + 2c_{\varphi Q}^{3}\,, \tag{2.7}$$

JHEP04(2019)100

as well as the combination of operators that affects the electromagnetic dipole of the top quark, defined as

$$c_{tA} \equiv \mathrm{Re}\{c_W C_{uB}^{(33)} + s_W C_{uW}^{(33)}\} = (c_{tW} - c_W c_{tZ})/s_W \,. \tag{2.8}$$

These two degrees of freedom, $c_{\varphi Q}^{+}$ and c_{tA}, are useful for instance in the interpretation of processes such as $Z \to b\bar{b}$ and $t\bar{t}\gamma$. Since they can be simply written as linear combinations of other degrees of freedom, we will not discuss them further in this work.

Taking stock, in total we have 9 CP-conserving degrees of freedom constructed from operators that involve two heavy quarks and gauge and Higgs bosonic fields. Operators involving gauge boson fields can be constrained either by single top production, if they modify the charged current coupling, or by the associated production of top quark pairs and single tops with electroweak bosons, i.e. processes such as $t\bar{t}V$ and tV, if they modify only the neutral current couplings. The degree of freedom c_{tG} will enter at leading-order in top pair and top pair associated production and tW, and at NLO in $t/s-$channel single top production. The degree of freedom $c_{t\varphi}$, on the other hand, can only be constrained from the associated production of a top quark pair with a Higgs boson, as we will discuss in the next section. Fortunately in this case, the first cross-section measurements for $t\bar{t}H$ production have recently become available.

Putting everything together, in total our fitting basis will be composed of $N_{\mathrm{op}} = 34$ independent degrees of freedom constructed from the dimension-6 SMEFT operators relevant for the description of the top quark sector: $N_{\mathrm{op}} = 11$ four-heavy-quark operators, $N_{\mathrm{op}} = 14$ two-heavy-two-light quark operators, and $N_{\mathrm{op}} = 9$ operators involving two heavy quarks and bosonic fields. In table 1 we summarise the definition of these 34 degrees of freedom in terms of the SMEFT operators in the Warsaw basis, as well as the internal notation that we will use in the following to refer to them. As in the above discussion, the degrees of freedom are divided into the three relevant classes: four-heavy-quark operators, two-heavy-two-light-quark operators, and operators that couple two heavy quarks to gauge and Higgs bosonic fields. We will discuss in the next section (see, in particular, table 6) which of these operators are constrained by each of the LHC top quark measurements included in the analysis.

Note that some of the degrees of freedom defined in table 1 have already been studied in the context of SMEFT fits of the top quark sector, see e.g. [48, 55, 56] and references therein. For instance, the chromomagnetic operator c_{tG} was constrained to be within $[-1.3, 1.2]$ for $\Lambda = 1\,\mathrm{TeV}$ at the 68% confidence level in the analysis of [48]. However, so far the simultaneous determination of the complete set of degrees of freedom of table 1 has never been carried out. In most cases, existing fits consider only either varying one operator at a time or marginalising over a smaller subset of operators, and the bounds derived in this way can differ significantly from those derived in a more global analysis. We will come back in section 5 to the comparison of the results of our analysis with previous studies in the literature.

The SMEFT degrees of freedom defined in table 1 have been implemented at LO in the UFO model `dim6top` as discussed in [10]. Results obtained with `dim6top` have been benchmarked with the independent UFO implementation available in the `SMEFTsim` package [57].

JHEP04(2019)100

Class	Notation	Degree of Freedom	Operator Definition
$QQQQ$	OQQ1	c^1_{QQ}	$2C^{1(3333)}_{qq} - \frac{2}{3}C^{3(3333)}_{qq}$
	OQQ8	c^8_{QQ}	$8C^{3(3333)}_{qq}$
	OQt1	c^1_{Qt}	$C^{1(3333)}_{qu}$
	OQt8	c^8_{Qt}	$C^{8(3333)}_{qu}$
	OQb1	c^1_{Qb}	$C^{1(3333)}_{qd}$
	OQb8	c^8_{Qb}	$C^{8(3333)}_{qd}$
	Ott1	c^1_{tt}	$C^{(3333)}_{uu}$
	Otb1	c^1_{tb}	$C^{1(3333)}_{ud}$
	Otb8	c^8_{tb}	$C^{8(3333)}_{ud}$
	OQtQb1	c^1_{QtQb}	$C^{1(3333)}_{quqd}$
	OQtQb8	c^8_{QtQb}	$C^{8(3333)}_{quqd}$
$QQqq$	O81qq	$c^{1,8}_{Qq}$	$C^{1(i33i)}_{qq} + 3C^{3(i33i)}_{qq}$
	O11qq	$c^{1,1}_{Qq}$	$C^{1(ii33)}_{qq} + \frac{1}{6}C^{1(i33i)}_{qq} + \frac{1}{2}C^{3(i33i)}_{qq}$
	O83qq	$c^{3,8}_{Qq}$	$C^{1(i33i)}_{qq} - C^{3(i33i)}_{qq}$
	O13qq	$c^{3,1}_{Qq}$	$C^{3(ii33)}_{qq} + \frac{1}{6}(C^{1(i33i)}_{qq} - C^{3(i33i)}_{qq})$
	O8qt	c^8_{tq}	$C^{8(ii33)}_{qu}$
	O1qt	c^1_{tq}	$C^{1(ii33)}_{qu}$
	O8ut	c^8_{tu}	$2C^{(i33i)}_{uu}$
	O1ut	c^1_{tu}	$C^{(ii33)}_{uu} + \frac{1}{3}C^{(i33i)}_{uu}$
	O8qu	c^8_{Qu}	$C^{8(33ii)}_{qu}$
	O1qu	c^1_{Qu}	$C^{1(33ii)}_{qu}$
	O8dt	c^8_{td}	$C^{8(33ii)}_{ud}$
	O1dt	c^1_{td}	$C^{1(33ii)}_{ud}$
	O8qd	c^8_{Qd}	$C^{8(33ii)}_{qd}$
	O1qd	c^1_{Qd}	$C^{1(33ii)}_{qd}$
$QQ + V, G, \varphi$	OtG	c_{tG}	$\mathrm{Re}\{C^{(33)}_{uG}\}$
	OtW	c_{tW}	$\mathrm{Re}\{C^{(33)}_{uW}\}$
	ObW	c_{bW}	$\mathrm{Re}\{C^{(33)}_{dW}\}$
	OtZ	c_{tZ}	$\mathrm{Re}\{-s_W C^{(33)}_{uB} + c_W C^{(33)}_{uW}\}$
	Off	$c_{\varphi tb}$	$\mathrm{Re}\{C^{(33)}_{\varphi ud}\}$
	Ofq3	$c^3_{\varphi Q}$	$C^{3(33)}_{\varphi q}$
	OpQM	$c^-_{\varphi Q}$	$C^{1(33)}_{\varphi q} - C^{3(33)}_{\varphi q}$
	Opt	$c_{\varphi t}$	$C^{(33)}_{\varphi u}$
	Otp	$c_{t\varphi}$	$\mathrm{Re}\{C^{(33)}_{u\varphi}\}$

Table 1. The notation that we will use to denote the results of the fits presented in this work. In each case, we indicate the internal notation for the degree of freedom and the corresponding definition in terms of the operators in the Warsaw basis. The degrees of freedom are divided into three classes: four-heavy-quark operators ($QQQQ$), two-heavy-two-light-quark operators ($QQqq$), and operators that couple two heavy quarks to gauge and Higgs bosonic fields ($QQ + V, G, \varphi$).

JHEP04(2019)100

In this work the `dim6top` model is complemented with the necessary counter-terms to enable NLO computations. The UFO model has been interfaced to `MadGraph5_aMC@NLO` to compute the $\mathcal{O}(\Lambda^{-2})$ and $\mathcal{O}(\Lambda^{-4})$ SMEFT corrections to the relevant SM cross-sections as indicated in eq. (2.2).

2.3 Operators not involving top quarks

In this work, we follow ref. [10] and only include those operators that explicitly modify the couplings of the top quark with the other SM fields. We therefore assume that other relevant operators are well constrained by processes that do not involve top quarks. In the following, we give a brief overview of these operators and discuss how they are constrained.

Firstly, the operator

$$O_G = f^{ABC} G_\mu^{A\nu} G_\nu^{B\rho} G_\rho^{C\mu}, \tag{2.9}$$

enters $t\bar{t}(V/H)$ production through a modification of the triple gluon coupling. Although it has been suggested that $t\bar{t}$ production can possibly constrain it due to its non-interference with the SM in di-jet production [58], it has been shown recently that this operator is most tightly constrained by multi-jet production measurements [59, 60]. The bounds on the coefficient c_G are found to lie within $[-0.04, 0.04]\,\text{TeV}^{-2}$, beyond the sensitivity of top quark pair production or associated production.

Secondly, operators involving a modification of the electroweak gauge-boson couplings to light fermions are in principle relevant to the interpretation of the single top, tZ and ttZ measurements. Most of these operators are however reasonably well constrained by electroweak precision observables. In the Warsaw basis, and under the assumed flavour structure, they are $\mathcal{O}_{\phi q}^{(1)}$, $\mathcal{O}_{\phi q}^{(3)}$, $\mathcal{O}_{\phi u}$, $\mathcal{O}_{\phi d}$, $\mathcal{O}_{\phi l}^{(3)}$, $\mathcal{O}_{\phi l}^{(1)}$, $\mathcal{O}_{\phi e}$, $\mathcal{O}_{ll}^{(3)}$, $\mathcal{O}_{\varphi WB}$ and $\mathcal{O}_{\varphi D}$. Among these 10 operators, 8 degrees of freedom are stringently constrained by electroweak observables [61], while two flat directions remain which are constrained only by diboson production processes, as discussed in refs. [53, 62, 63] for example. The two flat directions can be conveniently parametrised with [64]:

$$\mathcal{O}_{HW} = (D^\mu \phi)^\dagger \tau_I (D^\nu \phi) W_{\mu\nu}^I, \tag{2.10}$$

$$\mathcal{O}_{HB} = (D^\mu \phi)^\dagger (D^\nu \phi) B_{\mu\nu}, \tag{2.11}$$

which are linear combinations of Warsaw basis operators. Together with another basis operator O_W, they form the full set of operators that modify the triple-gauge-boson couplings (TGC). While in principle these couplings would enter the tZ process considered in this analysis (as well as measurements of $t\gamma$ that we do not include), they are well constrained from diboson production at LEP2 and the LHC. An interesting question is whether processes like tZ production could enhance the sensitivity to the anomalous TGCs, as the diagrams of this process in the SM display large cancellations among each other as required by unitarity, which are then spoiled by anomalous TGC leading to enhanced cross sections at large energy. The study performed in ref. [41] shows that while the effect is indeed present, it is not significant enough to compete with the sensitivities provided by diboson production. Therefore neglecting these operators in the associated production of single top quarks is well justified.

JHEP04(2019)100

Another operator that potentially would need to be taken into account is associated to the modification of the Zbb coupling, characterized by the coefficient of

$$O_{\varphi b} = (\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi)(\bar{b}\gamma^\mu b)\,, \tag{2.12}$$

as well as by the degree of freedom $c^+_{\varphi Q}$ defined in eq. (2.7). These are constrained by the decay rate of $Z \to b\bar{b}$ and the forward-backward asymmetry of $e^+e^- \to b\bar{b}$ at the Z pole measured by LEP. The corresponding constraints on the two coefficients are below $\mathcal{O}(0.1)\,\text{TeV}^{-2}$. Therefore $C_{\varphi b}$ can be safely ignored in this analysis. Conversely, the constraint on $c^+_{\varphi Q}$ is in principle relevant as it is a linear combination of $c^-_{\varphi Q}$ and $c^3_{\varphi Q}$, which enter this fit. However, in this work we choose not to include this information, because our goal is to quantify the direct constraints provided by top quark measurements.

Another operator that one might have to consider is the Higgs-gluon operator

$$O_{\varphi G} = \varphi^\dagger \varphi G^A_{\mu\nu} G^{A\mu\nu}\,, \tag{2.13}$$

which enters in the $t\bar{t}H$ production process. While this operator is already tightly constrained by Higgs production in gluon fusion, $gg \to H$, this process is also affected by exactly the same top quark degrees of freedom that enter $t\bar{t}H$, namely $c_{t\varphi}$ and c_{tG}. Therefore in principle the marginalised limit should be derived by combining $gg \to h$ together with the other top quark measurements, and fitting simultaneously $C_{\varphi G}$ and the relevant top quark operators. Such a combined fit to both the $t\bar{t}H$ and $gg \to H$ processes has been studied in [40], showing that, within its marginalised bound, $C_{\varphi G}$ does have an impact on the $t\bar{t}H$ rate, but it is not very significant. Therefore in this work we include the data on the $t\bar{t}H$ cross-sections but not the $gg \to H$ ones, since the latter will only fix the value of $C_{\varphi G}$ without affecting the description of the $t\bar{t}H$ process too much. We should keep in mind that it would be possible to improve upon this by explicitly including the $gg \to H$ cross-sections to our experimental inputs.

Another interesting operator is the following:

$$O_{bG} = \left(\bar{Q}\sigma^{\mu\nu} T^A b\right)\phi G^A_{\mu\nu}\,, \tag{2.14}$$

which, in addition to all four-fermion operators involving two light quarks and two right-handed b quarks, will affect the description of $t\bar{t}b\bar{b}$ production. Ref. [65] has reported $\mathcal{O}(1)\,\text{TeV}^{-2}$ bounds on C_{bG} from the analysis of $pp \to b\bar{b}$ production. The other four-fermion operators will enter the same process, and it is unlikely that $t\bar{t}b\bar{b}$ production will provide an even stronger constraint.

Finally, two-lepton-two-top-quark operators, such as $(\bar{t}\gamma^\mu t)(\bar{e}\gamma_\mu e)$ and $(\bar{Q}\gamma^\mu \tau^I Q)(\bar{l}\gamma_\mu \tau^I l)$, could in principle affect the description of $t\bar{t}Z$ and tZ production as well as the measurement of the W-helicity fractions in top quark decay, if the decays of W and Z bosons are taken into account. However, depending on the details of the analysis, the inclusion of these operators requires a reinterpretation of the experimental measurements, since for example the extrapolation from the fiducial to the total phase space could be affected by SMEFT effects. We thus postpone the inclusion of these operators to future studies.

JHEP04(2019)100

We emphasize that the decoupling of the operators that do not involve top quarks from the interpretation of top quark measurements at the LHC is in principle only an approximation. Not all of these operators are currently strictly constrained by other processes, but they could be dealt with either by possible improvements in the future, or by extending our fit by including additional measurements. However, until the most complete global fit can be performed, approximations like the ones we adopt here are always useful, because they allow us to focus on a certain sector of SMEFT, to study a certain type of processes, and to obtain an intuitive understanding of the underlying physics. Fortunately, as we can see from the discussions above, for a study focused on the interpretation of top quark measurements, the assumption of the decoupling of non-top operators is in general already very good, and future improvements can be envisioned. We therefore expect our results to be robust and not significantly affected by the possible inclusion of SMEFT operators that do not involve top quarks.

2.4 NLO QCD effects in the SMEFT calculation

In eq. (2.2), the coefficients κ_i and $\widetilde{\kappa}_{ij}$ can be evaluated at either leading order in both the QCD and electroweak couplings, or by also including higher-order perturbative corrections. Given the high precision of available top quark measurements, particularly from the LHC Run II, as well as the further improvements expected at Run II and during the High-Luminosity (HL) LHC, it is important to take into account the NLO QCD corrections to SMEFT effects. This is necessary for a number of reasons, including:

- QCD corrections to total rates are often quite large, especially for processes that are proportional to α_s at the Born level. Taking them into account results in general in an improvement of the bounds on the SMEFT Wilson coefficients. Additionally, NLO QCD corrections also reduce the theoretical uncertainties from scale variations, which is helpful in discriminating between different BSM scenarios.

- QCD corrections can distort the distributions of key observables. Given that the interpretation of differential distributions plays an important role in SMEFT global fits, providing reliable predictions for them is crucial. For instance, it is shown in ref. [66] that in the presence of a deviation from the SM, missing QCD corrections to certain differential distributions could lead us to make incorrect conclusions on the nature of BSM physics.

- The experimental sensitivity to SM deviations can be improved by using the most accurate SMEFT predictions and by optimizing the experimental strategies in a top-down way. However, the large QCD corrections at the LHC make this improvement unrealistic without consistently taking into account NLO predictions.

Motivated by the above considerations, theoretical calculations in the SMEFT at NLO in the QCD perturbative expansion have started to appear, in many cases matched with parton showers. For instance, computations including higher dimensional operators for

JHEP04(2019)100

top-related processes have been presented in [40, 41, 66–70], and those for Higgs and electroweak processes in [8, 71–76]. In particular, based on the machinery of automatic computations in `MadGraph5_aMC@NLO`, a systematic framework for including higher-dimensional operators at NLO has been established through the studies of refs. [40, 41, 66–70], and the implementation of the full set of dimension-6 SMEFT operators is currently being studied.

A novel feature of the present work, as compared to previous SMEFT studies of the top quark sector, is that we will exploit this framework and include the theoretical predictions at NLO in QCD whenever possible. This allows us to obtain the currently most accurate bounds on the coefficients of the SMEFT operators affecting the top quark couplings. Furthermore, by switching on and off the NLO QCD corrections in the fit, we can understand better the importance of the higher-order corrections in the SMEFT calculation when constraining different operators. In the following, we briefly explain which corrections will be included. As we will discuss in section 3, the SM calculations are always performed using the highest perturbative order available for each process.

Of all the degrees of freedom relevant in this work, the operators involving only two fermion fields have been fully automated already in this framework, it is therefore possible to straightforwardly evaluate their associated NLO QCD corrections. Four-fermion operators are being studied, and their complete implementation is expected to be publicly available in a short timescale. In this work we will include the NLO QCD corrections to the four-fermion operators only in the inclusive single top and top-pair production processes, which are the most accurately measured processes.

One practical difficulty in obtaining stable numerical results at NLO is that the residual uncertainties arising from the numerical Monte Carlo integration of the cross-sections, the so-called "MC errors", for the interference terms κ_i and in the cross terms $\widetilde{\kappa}_{ij}, i \neq j$ can be large. These κ terms are obtained by sampling the parameter space spanned by the full set of relevant degrees of freedom, and computing the total cross section (or other observables) iteratively. The results are fitted to the most general quadratic function. Therefore the MC error on interference/cross terms can be large, in particular when these terms are suppressed (see discussions in ref. [77]). Examples of suppressed interference contributions will be discussed in the next subsection. Given the large number of SMEFT operators relevant for the description of top quark measurements, a full simulation at NLO QCD would be very time consuming.

In this work, we adopt the following strategy:

- For $t\bar{t}$ and single top production, the experimental measurements exhibit the highest precision. We therefore use the full NLO simulation. This is done by sampling the parameter space following [77]. For each point, we generate 8×10^5 events, and estimate the corresponding MC errors for each observable included in the fit. These MC uncertainties can be taken into account when constructing the χ^2 function, as discussed in section 4.

- For associated production processes, the measurements are less accurate. We generate the full LO predictions using the implementation provided in [10]. We then apply K-factors from previous calculations of $t\bar{t}Z$, $t\bar{t}H$ and tZj production, wherever

JHEP04(2019)100

available [40, 41, 69]. For contributions or processes that have not been previously calculated (e.g. contributions from the four-fermion operators, and $t\bar{t}b\bar{b}$ and $tt\bar{t}\bar{t}$ processes), we simply apply the SM K-factor.

- The W-helicity in top quark decay is available at NLO in the form of analytical results [78].

In section 5 we will assess the stability of our results with respect to the inclusion or not of NLO QCD corrections to the SMEFT dimension-6 effects.

Finally, the $t\bar{t}t\bar{t}$ process is a special one, because the dominant contributions can come from $\mathcal{O}(c^4/\Lambda^8)$ terms, i.e. it goes beyond the parametrisation of eq. (2.2). This is due to diagrams with two insertions of $qqtt$ operators at the amplitude level, which leads to a rapid growth of the cross section as a function of energy, potentially causing problems with SMEFT validity [55]. In this work we only keep the terms up to $\mathcal{O}(\Lambda^{-4})$ from one insertion of the operators. This represents a good approximation if the coefficients of $qqtt$ operators are constrained to be within the order of a few TeV^{-2} at most.

Beyond this limit, $\mathcal{O}(\Lambda^{-8})$ terms dominate, and our predictions are not accurate, but they do give a lower bound of the true SMEFT contribution. Given that in practice, only the upper bound of this measurement is useful for constraining operators, our approximation will always lead to conservative result. In addition, to avoid possible EFT validity problems due to the particular energy-growth behavior of this process, we impose a hard cut of 2 TeV on the four-top invariant mass in our prediction. Once again this leads to a lower bound of the true SMEFT contribution. When compared with the upper bound of the full cross section measurement without the cut, it gives a conservative bound. In section 5.4 we will discuss the dependence of our final result on this cut.

2.5 General discussion

To complete this section, we briefly discuss some additional aspects of the SMEFT framework relevant to the present analysis of the top quark sector.

RG running and mixing. In general, the SMEFT operators $\{\mathcal{O}_i^{(6)}\}$ in eq. (2.1) will run with the scale and thus the coefficients $\{c_i\}$ will depend on the typical momentum transfer of the process. This dependence can be evaluated using renormalisation group (RG) equations [53, 79, 80], but here since we focus on processes with a similar energy scale, $E \simeq m_t$, we will not include these operator running effects. In any case, the inclusion of NLO QCD corrections will reduce this scale dependence [40, 81], even for the case of differential distributions where significantly different scales are involved. In [40], RGE effects are calculated for ttH production over a scale ranging from 150 GeV to 2 TeV. It is found that RGE effects remain below the 10% level. We expect a similar behaviour for other top production processes also at the differential level and therefore we can for this first study safely ignore RGE effects.

Energy enhancements. One important feature of eq. (2.2) is that certain SMEFT operators will induce a growth of the cross-sections $\sigma(E)$ with the energy E [4]. This

JHEP04(2019)100

is a consequence of the fact that, in four space-time dimensions, field theories involving operators with mass dimensions $d > 4$ exhibit a strong sensitivity to the UV cut-off of the theory Λ. In other words, the coefficients $\{\kappa_i\}$ in eq. (2.2) will include terms that grow quadratically with E. Restricting ourselves to the $\mathcal{O}(\Lambda^{-2})$ corrections, eq. (2.2) can be written schematically as:

$$\sigma(E) = \sigma_{\rm SM}(E)\left[1 + \sum_i^{N_{d6}} \omega_i \frac{c_i v^2}{\Lambda^2} + \sum_i^{N_{d6}} \widetilde{\omega}_i \frac{c_i E^2}{\Lambda^2} + \mathcal{O}(\Lambda^{-4})\right], \tag{2.15}$$

where v is the Higgs boson vacuum expectation value (vev). The coefficients $\{\omega_i\}$ and $\{\widetilde{\omega}_i\}$ are process-dependent and arise from the $\{\kappa_i\}$ coefficients in eq. (2.2), once we separate the energy-growing contributions.

Whether a dimension-6 operator leads to a nonzero value of the energy-growing coefficient $\tilde{\omega}_i$ depends on many factors. For example, four-fermion operators in $t\bar{t}$ and single top processes can interfere with the SM amplitude with the same helicity configurations without any additional suppression, and therefore their contributions are proportional to E^2/Λ^2 by simple power counting. On the contrary, operators involving only two fermions contribute as v^2/Λ^2: the current-current operators like $O_{\varphi Q}^{(1)}$ always enter with two powers of the Higgs vev, while the dipole operators come with one power of the Higgs vev, but the flip of the fermion chirality leads to an additional suppression factor m_t or m_b upon interfering with the SM.

The associated production channels are on the other hand more complicated. There, even two-fermion operators can lead to E^2/Λ^2 contributions. This is because already in the SM, each Feynman diagram could lead to energy-growing terms, but overall cancellations occur among the leading contributions in different diagrams as required by unitarity. With higher-dimensional operators instead, even a $\mathcal{O}(v^2/\Lambda^2)$ change in one diagram could spoil the cancellation and lead to $\mathcal{O}(E^2/\Lambda^2)$ modification of the total rate.

Finally, the specific observable under consideration also matters. The interference between SM and SMEFT amplitudes with different helicities are suppressed by mass factors, but this suppression can be lifted by considering the decay products of the particle with different helicities [82, 83]. Note that suppression due to different helicities only applies to the interference term. As a result, at large energy the SMEFT contribution could be dominated by the $\mathcal{O}(\Lambda^{-4})$ terms. This is actually one of the reasons to include the quadratic contributions from the operators. A similar situation occurs for instance in diboson production at the LHC, see the discussion in [84].

The schematic decomposition in eq. (2.15) indicates that the effects of those operators for which $\widetilde{\omega}_i \neq 0$ will be enhanced with the energy E of the process. In turn, this will lead to an increased sensitivity of the experimental measurements at high energies with the values of the corresponding c_i coefficients. This property can be uniquely exploited at the LHC, where multiple processes probe the TeV region. It has been shown that in some cases the enhancement due to energy dependence in LHC processes already leads to a sensitivity competitive with respect to the LEP measurements [85–87].

It is therefore an interesting question whether the same happens for top quark measurements, and which operators can benefit from the high energy reach provided by the

JHEP04(2019)100

LHC. In this work we are going to study the impact of high energy measurements by comparing results with and without high mass bins in differential distributions, for instance in the invariant mass distribution in $t\bar{t}$ production. This will be discussed in section 5.4.

Validity of the SMEFT. Following up on the previous discussion, while it is useful to make use of the large energy transfer in the observed events, one has to pay special attention to remain in the region $E \ll \Lambda$. Otherwise, the whole validity of the SMEFT power expansion would be questionable, and it would become impossible to interpret the resulting constraints within any explicit BSM model.

To make sure that SMEFT analyses remain in their validity region, it has been proposed in ref. [88] and recommended in the Top LHC WG EFT note [10] that a kinematic cut E_{cut} should be imposed to events being analysed, as an upper bound of the energy transfer, so that the condition

$$E < E_{\text{cut}} < \Lambda \tag{2.16}$$

is always guaranteed. Given that Λ is the scale of the BSM dynamics and is model-dependent, different values of E_{cut} should be used, and results should be derived for each of these values. While the input data used in this work is not provided with such explicit cut, for specific distributions it is possible to remove bins with scale higher than a given value of E_{cut}. A strong dependence of the final results on high mass bins would imply that the sensitivity is dominated by the high energy events, and that the constraints can only be interpreted for those BSM models where Λ lies above the scale of the largest bin used in the fit.

Among the input data used in this work, only the $m_{t\bar{t}}$ distribution in $t\bar{t}$ production extends to energy scales above 1 TeV. Therefore, in section 5.4 we will study the dependence of our results on the high mass bins in the $m_{t\bar{t}}$ distributions of ATLAS and CMS.

Quadratic dimension-6 contributions. The last term in eq. (2.2) arises from the squares of dimension-six amplitudes. At order $\mathcal{O}(\Lambda^{-4})$, they are formally subleading contributions, so one can decide to include them without modifying the accuracy of the prediction of the central value. Indeed, there are good reasons why it could be worth including them in the analysis. To begin with, in BSM models with relatively large couplings the quadratic dimension-6 terms can become dominant without exiting the realm of validity of the EFT, see for instance refs. [88–91]. Therefore including the quadratic terms allows the results of a SMEFT analysis to be interpreted in the context of these scenarios.

Furthermore, the interference terms in eq. (2.2) are often suppressed, so that the leading contributions arise from quadratic dimension-6 terms. In these cases, one relies on the quadratic terms to extract meaningful bounds from the measurements. As an extreme case, the SM amplitude may not interfere with the SMEFT amplitude at all, because of different helicity and colour structures or different CP parity. As an illustration, in our analysis $c_{\varphi tb}$ and c_{bW} cannot interfere with the SM in the limit of $m_b \to 0$, so they can only be constrained once $\mathcal{O}(\Lambda^{-4})$ terms are included in the fit. Similarly, several of the $qqtt$ operators do not interfere with SM due to their color singlet interaction, though this is slightly lifted once NLO corrections are added.

JHEP04(2019)100

It is also possible that, while the interference term exists, it does not lead to an energy growth behaviour. We have already mentioned that some operators cannot lead to a nonzero $\tilde{\omega}_i$ at the order $\mathcal{O}(\Lambda^{-2})$. One would then expect that the dominant sensitivity for these operators comes from $\mathcal{O}(\Lambda^{-4})$ contributions at large energies. For instance, it has been observed that in diboson production at the LHC, the helicity selection rule [92] leads to an energy suppression in the interference term, so the sensitivity to TGC couplings is dominated by $\mathcal{O}(\Lambda^{-4})$ terms [84]. Finally, a suppression of the interference term could be simply accidental. This has been observed for weak dipole operators in $t\bar{t}Z$ and $t\bar{t}\gamma$ production processes [69], and is relevant also in the present analysis.

Here we will follow the recommendations of ref. [10] and repeat the analysis with and without including the quadratic SMEFT contributions. This comparison will then tell us where quadratic dimension-6 contributions are subleading and where the truncation at dimension-6 at the cross section level can be a good approximation. In addition, under certain assumptions, the $\mathcal{O}(\Lambda^{-4})$ corrections could also provide an estimate of how reliable are the SMEFT fit results with respect to higher orders in the effective theory parameter expansion. Note, however, that for the degrees of freedom whose contributions at $\mathcal{O}(\Lambda^{-2})$ are extremely suppressed, such as $c_{\varphi tb}$ and c_{bW}, and for the color-singlet four-fermion interactions $c_{Qq}^{1,1}$, c_{tu}^{1}, c_{td}^{1}, c_{tq}^{1}, c_{Qu}^{1}, and c_{Qd}^{1}, our numerical approach would lead to large MC errors on the interference, and therefore the resulting bounds from a $\mathcal{O}(\Lambda^{-2})$ fit will be at most qualitative.

3 Experimental data and theoretical calculations

In this section we describe the experimental measurements of top quark production at the LHC which will be used to constrain the SMEFT operators related to the top sector. For each dataset we discuss its main features, the information that it provides on the SMEFT effects, and the treatment of experimental uncertainties. We also describe the settings of the theoretical calculations of the SM and SMEFT contributions to the cross-sections that are used for each process. Finally, we summarise the main features of our choice of fitting basis in terms of the sensitivity of each of the input LHC processes to the individual operators.

3.1 Top quark production at the LHC

In the present analysis, we will constrain the top quark sector of the SMEFT by using experimental measurements from the LHC Run I at $\sqrt{s} = 8\,\text{TeV}$ and from Run II at $\sqrt{s} = 13\,\text{TeV}$. We do not consider previous, less precise data at $\sqrt{s} = 7\,\text{TeV}$ nor data from the Tevatron. The measurements of top quark production at the LHC 7 TeV are superseded by the more precise 8 and 13 TeV ones and there is no loss of information incurred by not including them. Concerning the Tevatron measurements, on the one hand, they are affected by larger uncertainties than those of the LHC, but on the other hand the production of top quarks proceeds mostly via the quark-antiquark process and therefore it provides additional sensitivity to new combinations of four-fermion operators. While in this work we have focused on LHC data, we plan to include the Tevatron constraints in a future iteration of our analysis.

JHEP04(2019)100

Process	Dataset	$\sqrt{s}$	Info	Observables	$N_{\rm dat}$	Ref.
$t\bar{t}$	ATLAS_tt_8TeV_ljets	**8 TeV**	lepton+jets	$d\sigma/d\|y_t\|$, $d\sigma/dp_t^T$, $d\sigma/dm_{t\bar{t}}$, $d\sigma/d\|y_{t\bar{t}}\|$	5, 8, 7, 5	[93]
$t\bar{t}$	CMS_tt_8TeV_ljets	**8 TeV**	lepton+jets	$d\sigma/dy_t$, $d\sigma/dp_t^T$, $d\sigma/dm_{t\bar{t}}$, $d\sigma/dy_{t\bar{t}}$	10, 8, 7, 10	[94]
$t\bar{t}$	CMS_tt2D_8TeV_dilep	**8 TeV**	dileptons	$d^2\sigma/dy_t dp_t^T$, $d^2\sigma/dy_t dm_{t\bar{t}}$, $d^2\sigma/dp_{t\bar{t}}^T dm_{t\bar{t}}$, $d^2\sigma/dy_{t\bar{t}} dm_{t\bar{t}}$	16, 16, 16, 16	[95]
$t\bar{t}$	CMS_tt_13TeV_ljets	**13 TeV**	lepton+jets	$d\sigma/d\|y_t\|$, $d\sigma/dp_t^T$, $d\sigma/dm_{t\bar{t}}$, $d\sigma/d\|y_{t\bar{t}}\|$	7, 9, 8, 6	[98]
$t\bar{t}$	CMS_tt_13TeV_ljets2	**13 TeV**	lepton+jets	$d\sigma/d\|y_t\|$, $d\sigma/dp_t^T$, $d\sigma/dm_{t\bar{t}}$, $d\sigma/d\|y_{t\bar{t}}\|$	11, 12, 10, 10	[100]
$t\bar{t}$	CMS_tt_13TeV_dilep	**13 TeV**	dileptons	$d\sigma/dy_t$, $d\sigma/dp_t^T$, $d\sigma/dm_{t\bar{t}}$, $d\sigma/dy_{t\bar{t}}$	8, 6, 6, 8	[101]
$t\bar{t}$	ATLAS_WhelF_8TeV	**8 TeV**	W helicity fract	F_0, F_L, F_R	3	[96]
$t\bar{t}$	CMS_WhelF_8TeV	**8 TeV**	W helicity fract	F_0, F_L, F_R	3	[97]

Table 2. The experimental measurements of inclusive top quark pair production at the LHC considered in the present analysis to constrain the coefficients of the SMEFT dimension-6 operators in the top sector. For each dataset, we indicate the type of process, the dataset label, the center of mass energy $\sqrt{s}$, the final state or the specific production mechanism, the available observables, the number of data points $N_{\rm dat}$, and the publication reference. Most distributions are statistically correlated among them and one needs to be careful to avoid double counting.

The various experimental datasets used as input in this work are summarised in tables 2, 3 and 4, for inclusive top quark pair production, $t\bar{t}$ production in association with gauge and Higgs bosons, and single top production measurements, respectively. For each dataset, we indicate the type of process, its label, the centre-of-mass energy $\sqrt{s}$, information on the final state or the specific production mechanism, the available observables, the number of data points $N_{\rm dat}$, and the corresponding publication reference.

As we will discuss in more detail below, information on correlations between systematic uncertainties is available only for a subset of the data, specifically, for all the 8 TeV distributions in table 2 (including W helicity fractions) [93–97] and for the top quark pair production measurement at 13 TeV of ref. [98]. For the rest of the datasets, since this information is missing, we add statistical and systematic uncertainties in quadrature. Also, the correlation matrix among various differential distributions is not usually available. To avoid double counting, only one distribution per dataset can therefore be included in the fit. The ATLAS 8 TeV lepton+jet dataset is an exception, where such correlations have recently become available [99]. However, because the effect of correlating all differential distributions has not been studied yet in the fitting framework used here, we do not utilise them.

As indicated in table 2, in the case of differential distributions we always use absolute rather than normalised cross-sections. The rationale is that absolute distributions are more sensitive to SMEFT effects than the normalised ones, except when the corresponding

JHEP04(2019)100

Process	Dataset	$\sqrt{s}$	Info	Observables	N_{dat}	Ref.
$t\bar{t}b\bar{b}$	CMS_ttbb_13TeV	**13 TeV**	total xsec	$\sigma_{\text{tot}}(t\bar{t}b\bar{b})$	1	[102]
$t\bar{t}t\bar{t}$	CMS_tttt_13TeV	**13 TeV**	total xsec	$\sigma_{\text{tot}}(t\bar{t}t\bar{t})$	1	[103]
$t\bar{t}Z$	CMS_ttZ_8_13TeV	**8+13 TeV**	total xsec	$\sigma_{\text{tot}}(t\bar{t}Z)$	2	[104, 105]
$t\bar{t}Z$	ATLAS_ttZ_8_13TeV	**8+13 TeV**	total xsec	$\sigma_{\text{tot}}(t\bar{t}Z)$	2	[106, 107]
$t\bar{t}W$	CMS_ttW_8_13TeV	**8+13 TeV**	total xsec	$\sigma_{\text{tot}}(t\bar{t}W)$	2	[104, 105]
$t\bar{t}W$	ATLAS_ttW_8_13TeV	**8+13 TeV**	total xsec	$\sigma_{\text{tot}}(t\bar{t}W)$	2	[106, 107]
$t\bar{t}H$	CMS_tth_13TeV	**13 TeV**	signal strength	$\mu_{t\bar{t}H}$	1	[108]
$t\bar{t}H$	ATLAS_tth_13TeV	**13 TeV**	total xsec	$\sigma_{\text{tot}}(t\bar{t}H)$	1	[109]

Table 3. Same as table 2 now for $t\bar{t}$ in association with heavy quarks, with weak vector bosons, and with the Higgs boson.

Process	Dataset	$\sqrt{s}$	Info	Observables	N_{dat}	Ref.
Single t	CMS_t_tch_8TeV_inc	**8 TeV**	t-channel	$\sigma_{\text{tot}}(t), \sigma_{\text{tot}}(\bar{t})$ (R_t)	2 (1)	[110]
Single t	CMS_t_sch_8TeV	**8 TeV**	s-channel	$\sigma_{\text{tot}}(t+\bar{t})$	1	[111]
Single t	ATLAS_t_sch_8TeV	**8 TeV**	s-channel	$\sigma_{\text{tot}}(t+\bar{t})$	1	[112]
Single t	ATLAS_t_tch_8TeV	**8 TeV**	t-channel	$d\sigma(tq)/dp_T^t, d\sigma(\bar{t}q)/dp_T^{\bar{t}}$ $d\sigma(tq)/dy_t, d\sigma(\bar{t}q)/dy_t$	5, 4 4, 4	[113]
Single t	ATLAS_t_tch_13TeV	**13 TeV**	t-channel	$\sigma_{\text{tot}}(t), \sigma_{\text{tot}}(\bar{t})$ (R_t)	2 (1)	[114]
Single t	CMS_t_tch_13TeV_inc	**13 TeV**	t-channel	$\sigma_{\text{tot}}(t+\bar{t})$ (R_t)	1 (1)	[115]
Single t	CMS_t_tch_8TeV_dif	**8 TeV**	t-channel	$d\sigma/dp_T^{(t+\bar{t})}$, $d\sigma/d\lvert y^{(t+\bar{t})}\rvert$	6 6	[116]
Single t	CMS_t_tch_13TeV_dif	**13 TeV**	t-channel	$d\sigma/dp_T^{(t+\bar{t})}$, $d\sigma/d\lvert y^{(t+\bar{t})}\rvert$	4 4	[117]
tW	ATLAS_tW_inc_8TeV	**8 TeV**	inclusive	$\sigma_{\text{tot}}(tW)$	1	[118]
tW	CMS_tW_inc_8TeV	**8 TeV**	inclusive	$\sigma_{\text{tot}}(tW)$	1	[119]
tW	ATLAS_tW_inc_13TeV	**13 TeV**	inclusive	$\sigma_{\text{tot}}(tW)$	1	[120]
tW	CMS_tW_inc_13TeV	**13 TeV**	inclusive	$\sigma_{\text{tot}}(tW)$	1	[121]
tZ	CMS_tZ_inc_13TeV	**13 TeV**	inclusive	$\sigma_{\text{fid}}(Wbl^+l^-q)$	1	[122]
tZ	ATLAS_tZ_inc_13TeV	**13 TeV**	inclusive	$\sigma_{\text{tot}}(tZq)$	1	[123]

Table 4. Same as table 2, now for inclusive single t production and single t production in association with vector bosons.

fiducial cross-sections are included in the fit at the same time. If the measurements are presented only in terms of normalised differential distributions, absolute distributions are reconstructed from the former using the fiducial cross-section. Uncertainties are added in quadrature. To avoid double counting, total and/or fiducial cross-sections are excluded from the fit whenever the corresponding absolute differential distributions are part of the input dataset.

JHEP04(2019)100

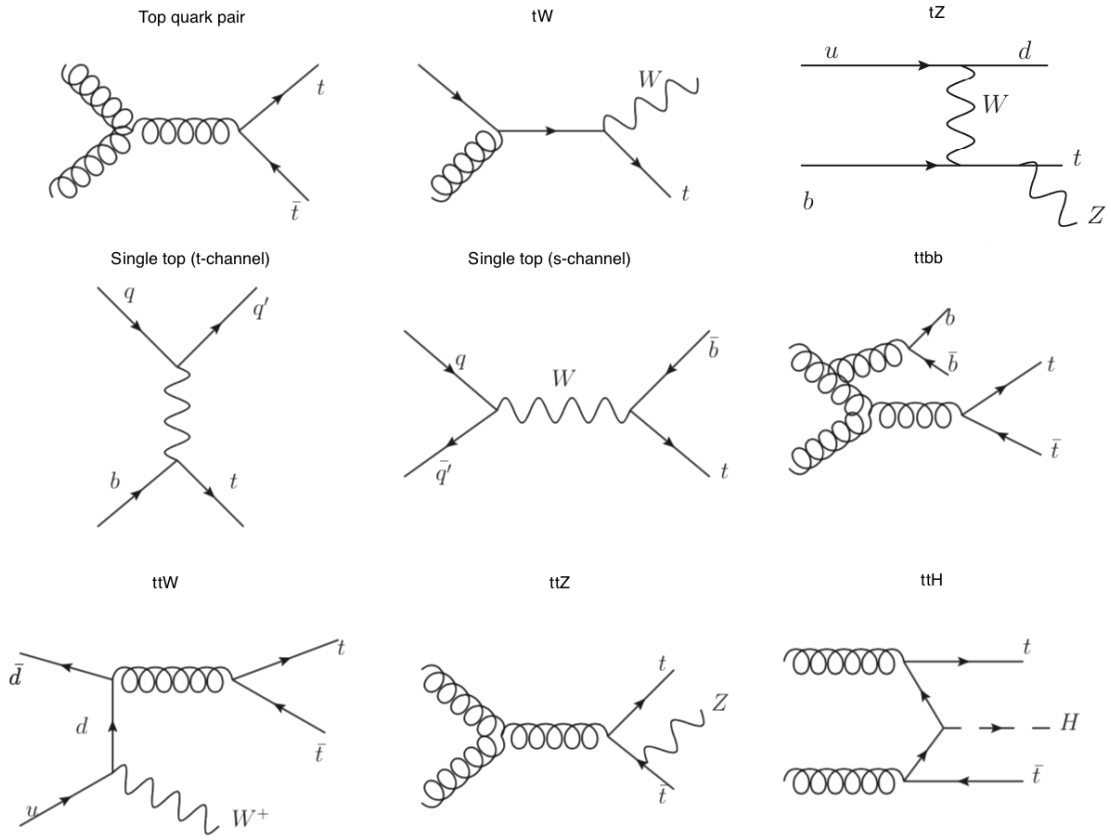

Figure 1. Representative Feynman diagrams at the Born level for the dominant production channels of top quarks at the LHC that are considered in the present SMEFT analysis. We show top-quark pair production; single-top production in association with a W or Z boson and in the t- and s-channels; $t\bar{t}$ production in association with $t\bar{t}$ or $b\bar{b}$; and $t\bar{t}$ production in association with a W or Z gauge boson or with the Higgs boson H.

To gain some intuition about the expected sensitivity of the input dataset to each of the SMEFT operators defined in section 2, it is useful to recall what are the dominant production mechanism for each top-related observable in the SM. In figure 1 we display representative Feynman diagrams at the Born level for the production of top quarks at the LHC in the channels that we consider in this analysis. Specifically, we show top-quark pair production; single-top production in association with a W or Z boson and in the t- and s-channels; $t\bar{t}$ production in association with $t\bar{t}$ or $b\bar{b}$; and $t\bar{t}$ production in association with a W or Z gauge boson or with the Higgs boson H.

From these diagrams, one can see that measurements of inclusive top quark pair production will be particularly sensitive to SMEFT operators that induce or modify interactions of the form $gt\bar{t}$ and $ggt\bar{t}$, such as the chromomagnetic operator c_{tG}. In this case, the interference with the most relevant SM production mechanism will dominate over the small quark-antiquark-initiated contributions. Likewise, single-top production and associated tW and tZ production will constrain SMEFT operators that involve both top quarks and electroweak gauge bosons, such as c_{tW}. As a third example, $tt\bar{b}\bar{b}$ production should provide direct information on operators involving four heavy quarks, such as c_{QQ}^1 and c_{QQ}^8.

JHEP04(2019)100

Following this overview of our input dataset, we move to describe in more detail the features of the individual measurements listed in tables 2, 3 and 4.

3.2 Top quark pair production

We begin by presenting the LHC datasets of top quark pair production used in this work. We consider inclusive production first, and then $t\bar{t}$ production in association with heavy quarks, with an electroweak gauge boson, and with the Higgs boson.

Inclusive top-quark pair production. At the LHC, the dominant mechanism for the production of top quarks is through the production of $t\bar{t}$ pairs. The inclusive $t\bar{t}$ process is dominated by the gluon-gluon initial state, with a small admixture of the quark-antiquark partonic luminosity [124]. In this analysis, we will limit ourselves to parton-level distributions constructed in terms of the kinematical variables of the top and anti-top quark, for which NNLO QCD corrections are available in the SM [125]. See [126] for recent progress in higher order calculations at the particle level for decayed top quarks, in terms of leptons and b-jets. For all the inclusive $t\bar{t}$ processes computed here, the SM prediction is computed up to NNLO in the QCD coupling. Theoretical predictions are obtained at NLO with `Sherpa` [127], for 8 TeV measurements, and with `MCFM` [128], for 13 TeV measurements, and are then supplemented with the NNLO QCD K-factors computed in ref. [129].

In the present analysis we include the ATLAS and CMS differential distributions from $t\bar{t}$ production at $\sqrt{s} = 8$ TeV in the lepton+jets final state [93, 94]. These measurements are those used in the study of [130] to constrain the large-x gluon PDF from the $t\bar{t}$ differential cross-sections, and are part of the NNPDF3.1 input dataset [36]. In both cases, the distributions in top quark transverse momentum and rapidity, p_T^t and y_t, as well as in top-quark pair invariant mass and rapidity, $m_{t\bar{t}}$ and $y_{t\bar{t}}$ are available, both as absolute cross-sections and normalised to the inclusive results; only the former are used here. As discussed in [130], to avoid double counting only one distribution per experiment can be added to the fit, as long as correlations between different distributions are not available or neglected.

Besides these two datasets, we take into account the constraints from the double-differential distributions from CMS at 8 TeV, which provide a good handle on the underlying partonic kinematics [95]. Note that this dataset is based on the dilepton final state, therefore it does not overlap with the dataset used in [94], which instead is based on the lepton+jets channel. We also include the CMS differential distributions at $\sqrt{s} = 13$ TeV in the lepton+jets [98] and dilepton [101] final states based on an integrated luminosity of $\mathcal{L} = 2.3\,\text{fb}^{-1}$, as well as the more recent measurements in the lepton+jet channel based on $\mathcal{L} = 35.8\,\text{fb}^{-1}$ [100]. A measurement based on the same dataset but with the dilepton final state was presented in [131]. Double-differential distributions from CMS at 13 TeV [98] are excluded since they overlap with the single-inclusive distributions from the lepton+jets datasets.

We do not include ATLAS measurements at 13 TeV since the published differential cross-sections at 13 TeV in the lepton+jets [132] and dilepton [133] channels are provided at the particle level. In this work, we restrict ourselves to parton-level observables. Note that in principle ATLAS measurements at 13 TeV are also available for the fully hadronic

JHEP04(2019)100

final state in the highly boosted regime [134]. These measurements are not considered here since their analysis requires jet substructure information alongside the consistent inclusion of electroweak [135] and threshold resummation [136] corrections.

Helicity fractions and spin correlations in $t\bar{t}$ production. A further window on the underlying dynamics of top quark pair production is provided by the measurement of observables sensitive to the spin structure of top quark production and decay. Among them, polarisation, W helicity fractions, and spin correlations provide direct constraints on the structure of the tWb vertex. In this work, we include the helicity fractions F_L, F_0, and F_R of the W bosons from the decay of top quarks measured by ATLAS [96] and CMS [97] at $\sqrt{s} = 8\,\text{TeV}$. These measurements supersede the previous ATLAS and CMS combined analysis at $\sqrt{s} = 7\,\text{TeV}$ [137].

Related types of available angular observables in $t\bar{t}$ production include the polarisation asymmetry $A_{P^\pm}$, the spin correlations variable $A_{\Delta\phi}$, and the $A_{c_1c_2}$ and $A_{\cos\phi}$ asymmetries, which discriminate between the correlated and uncorrelated t and $\bar{t}$ spins. For example, CMS has presented measurements of $t\bar{t}$ spin correlations and top quark polarisation in the lepton+jet and dilepton final states at $\sqrt{s} = 8\,\text{TeV}$ [138]. Measurements of the $t\bar{t}$ spin correlations at $\sqrt{s} = 13\,\text{TeV}$ in the $e\mu$ final state are also available from ATLAS, specifically the differential cross-section in the angular separation between the two leptons.This measurement deviates from the SM predictions by more than three sigma [139]. We leave the inclusion of these observables in the SMEFT global fit to future work.

$t\bar{t}V$ production. In this analysis we also include data for the production of a $t\bar{t}$ pair in association with either a Z or a W boson, which is directly sensitive to the top quark couplings with the gauge bosons (see figure 1). Specifically, we include the measurements of the total inclusive cross-sections for $t\bar{t}Z$ and $t\bar{t}W$ production at $\sqrt{s} = 8\,\text{TeV}$ and $\sqrt{s} = 13\,\text{TeV}$ from ATLAS [106, 107] and CMS [104, 105]. Note that, for $t\bar{t}W$, the W boson is often emitted from initial-state light quarks, however, when it is emitted from a final-state leg, it becomes sensitive to operators involving only one heavy quark, which is a unique feature of this process. We do not include the $t\bar{t}\gamma$ production measurements [140–142], whose interpretation is hampered by issues related to photon isolation and fragmentation, as well as to initial- and final-state radiation. Because of electroweak symmetry, the $t\bar{t}V$ process is closely related to the $t\bar{t}H$ one, to be discussed next.

Higgs production in association with a $t\bar{t}$ pair. The production of a top-antitop pair together with a Higgs boson allows for a direct probe of the Yukawa coupling of the top quark, as illustrated by the dominant mechanism indicated in figure 1. Recently, 5σ evidence for this production mode was presented by both the ATLAS and CMS collaborations [108, 109]. In the CMS case [108], we utilise their measurement of the signal strength $\mu_{t\bar{t}h}$ at $\sqrt{s} = 13\,\text{TeV}$ (normalised to the SM prediction), rather than the cross-section, because the latter is obtained by combining data at different centre-of-mass energies. In the ATLAS case [109], we utilise their measurement of the total cross-section for $t\bar{t}h$ production at $\sqrt{s} = 13$, extrapolated to the full phase space.

JHEP04(2019)100

***$t\bar{t}b\bar{b}$* and *$t\bar{t}t\bar{t}$* production.** The production of a top quark pair in association with a bottom-antibottom pair is a purely QCD process, where a $b\bar{b}$ pair is radiated either from a gluon emitted from the initial state or from the final state (see figure 1). The production of four top quarks at the LHC, $t\bar{t}t\bar{t}$, obeys a similar underlying mechanism in the SM, with the cross-section now being rather smaller due to the heavier top quark mass. The relevance of this process on the top quark sector of the SMEFT has been discussed in ref. [45].

Concerning $t\bar{t}b\bar{b}$, the total cross-section for $t\bar{t}b\bar{b}$, production, extrapolated to the full phase space at $\sqrt{13}\,\text{TeV}$, is available from CMS [102], together with the corresponding ratio to $t\bar{t}jj$ production. This single data point is included in our fit. Differential cross-sections for $t\bar{t}b\bar{b}$ production at $\sqrt{s} = 8\,\text{TeV}$ have also been presented as a function of the kinematics of the b-jets [143]. These measurements, however, are not included in the fit. While there is a priori no reason why these measurements of $t\bar{t}$ production in association with jets cannot be used in a SMEFT fit [144], this would require to interface the parton level calculation to a Monte Carlo shower program and possibly merge samples of different jet multiplicity in order to appropriately model the extra hadronic radiation. In this work, we have chosen to focus on inclusive observables, and we have deferred more exclusive observables to future work.

The ATLAS collaboration has presented the results of the measurements of top quark pair production in association with multiple b-jets at $\sqrt{s} = 13\,\text{TeV}$. Fiducial cross-sections for tt in association with more than two b quarks, both in the lepton+jets and in the dilepton channels, are provided [145], which supersede a previous measurement at 8 TeV [146]. We do not include these results in the fit as the cuts needed to simulate the cross-sections are not fully provided, and are therefore not reproducible.

Concerning $t\bar{t}t\bar{t}$, a first measurement of its cross-section at $\sqrt{s} = 13\,\text{TeV}$ has been presented by CMS [103], albeit with a statistical significance of only 1.6σ. This measurement supersedes previous upper bounds at 8 TeV [147] and 13 TeV [148]. In the case of ATLAS, upper bounds based on the 2015 dataset at 13 TeV were presented in [149] and then updated in [150] from the 2016 dataset. In this analysis we utilise the CMS cross-section measurement of [103].

3.3 Single top quark production

We turn to discuss single top quark production, first inclusively in either the t- or the s-channel, and then in association with an electroweak gauge boson.

Inclusive single top quark production. As highlighted in table 1, some of the SMEFT $d = 6$ operators that contribute to single top production via interference with the SM amplitudes are different from the corresponding ones in top quark pair production, whence their relevance in a global fit. There exist three main modes to produce single top quarks [151]: by means of the exchange of a $W^\pm$ boson, either in the t-channel or in the s-channel, and by means of the associated production with a $W^\pm$ (Z) boson that leads to the $q't$, $t\bar{b}$, and $tW^\pm$ (tZq) Born-level final states. Representative diagrams for these three modes are shown in figure 1.

JHEP04(2019)100

In this work, we include all relevant single top production datasets in the t- and s-channels from ATLAS and CMS at 8 and 13 TeV, see table 4. We restrict ourselves to parton-level measurements, that is, to un-decayed top quarks.

From ATLAS, we include the differential cross-sections at $\sqrt{s} = 8$ TeV [114], specifically the $d\sigma(tq)/dp_T^t$ and $d\sigma(tq)/dy_T^t$ distributions, as well as the corresponding measurements for anti-top quarks. From CMS, we include the inclusive cross-sections for t and for $\bar{t}$ production at $\sqrt{s} = 8$ TeV [110], as well as the corresponding differential distributions in $p_T^{(t+\bar{t})}$ and $|y_{t+\bar{t}}|$ [116]. In the case of the inclusive measurements, the ratio $R_t = \sigma(tq)/\sigma(\bar{t}q)$ is also provided, the use of which would be advantageous if the knowledge on correlations were lacking, due to the partial cancellation of experimental and theoretical systematic uncertainties between the numerator and the denominator.

We now move to single top t-channel based on the Run II dataset at $\sqrt{s} = 13$ TeV. We include the transverse momentum $p_T^{t+\bar{t}}$ and rapidity $|y_{t+\bar{t}}|$ differential distributions for single top production from CMS [117], the ATLAS and CMS measurements of the total inclusive cross-sections for single t and $\bar{t}$ production [114, 115]. The ratio R_t is once more provided in both cases.

Concerning single top s-channel measurements, we include the CMS total cross-sections in the s-channel at 8 TeV [111]. We also include the total cross-sections at 8 TeV from ATLAS [112]. No measurements of s-channel single top production at 13 TeV are available from either experiment. Neither the ATLAS nor the CMS differential distributions are provided with a full breakdown of experimental systematic uncertainties. Therefore, we sum all statistical and systematic uncertainties in quadrature. To avoid double counting, we do not include total cross-sections if the corresponding absolute differential distributions are already part of the input dataset. For example, using the labelling of table 4, if the `CMS_t_tch_13TeV_dif` distributions are used, then the associated `CMS_t_tch_13TeV_inc` total cross-sections are excluded from the fit.

For both inclusive and differential single-top measurements in the t-channel, NNLO QCD corrections have been computed [152]. For all single-top processes for which the measurements have been published (all measurements except the CMS differential measurements), we use as a theory input the NNLO calculation. For unpublished measurements, we use NLO QCD, as the NNLO results are not available. For the calculation of single top and single top associated channels at NLO, we use a $n_f = 5$ scheme for consistency for both the SM and the SMEFT calculations. Given that we are using NLO calculations for all single top and single top associated production processes, we expect the uncertainties related to the choice of number flavour scheme to be under control.

***tV* associated production.** The associated production of a top quark and a W boson has a very distinctive signature that allows one to reconstruct the decay products from both the top quark and the W decay. Measurements of tW associated production have been presented by ATLAS and CMS. Here we include the ATLAS measurements at 8 and 13 TeV of the total $\sigma(tW)$ cross-section [118, 120] extrapolated to the full phase space. A measurement of differential distributions at 13 TeV based on a luminosity of $\mathcal{L} = 36\,\text{fb}^{-1}$ [153] was also presented by ATLAS. However, this measurement is at particle level (of leptons

JHEP04(2019)100

and b-jets from the W and top quark decays), therefore we do not include it in the fit. We also include the CMS measurements of $\sigma_{\rm tot}(tW)$ at 8 and 13 TeV [119, 121]. These measurements supersede the previous ones at 7 and 8 TeV [119, 154].

For tW production both the SM and SMEFT part is for consistency computed with the DR1 scheme [155]. We refer the reader to ref. [156] for a detailed discussion of the differences between the various diagram subtraction/removal schemes. At the inclusive level the difference between the schemes is about 10% but this is greatly reduced with fiducial cuts. This uncertainty is comparable to the scale and PDF uncertainties and we do not expect it to significantly alter our findings.

The associated production of a single top quark in association with a Z boson, shown in figure 1, is also an interesting probe of the top quark sector of the SMEFT. The tZ production cross-section has been measured by CMS at 13 TeV in the Wbl^+l^-q final state, where the dilepton pair arise from the decay of the Z boson [122] (see [157] for an update based on $\mathcal{L} = 77.4\,{\rm fb}^{-1}$). The tZ production cross-section has been measured at 13 TeV by ATLAS in the tri-lepton final state and extrapolated to the full phase space [123]. We use these two measurements as data points in the fit, for a total of four tV input cross-sections.

3.4 Theory overview and sensitivity to the SMEFT degrees of freedom

In table 5, we summarise the details of the theoretical calculations used for the description of the LHC top quark production measurements included in the present analysis. We indicate, for both the SM and the SMEFT contributions to the cross-sections in eq. (2.2), the perturbative accuracy and the codes used to produce the corresponding predictions. In all cases, the same theoretical settings have been used for the calculation of both the total cross-sections and the differential distributions, where available. We emphasise that we have used state-of-the-art theory calculations for both the SM and the SMEFT pieces, which are instrumental to reduce the theoretical uncertainties associated to the missing perturbative higher orders. We have adopted common input settings for the theory calculations, in particular all the SM cross-sections are consistently evaluated with a NNPDF3.1 PDF set at NNLO accuracy that does not include any top data (henceforth labelled NNPDF3.1NNLO no-top). We do not include top quark data in the PDF fit as the datasets used in NNPDF are also used in the current SMEFT fit. Adding $t\bar{t}$ distributions into both fits would imply to double-counting the data, otherwise, as MC replicas are also used in NNPDF it would require us to keep track of correlations between the two sets of MC replicas. We therefore choose to exclude the top quark data from the PDF fit.

It should be clear from the above discussion, as well as from the considerations presented in section 2, that each of the input LHC processes will have a rather different sensitivity to each of the $N_{\rm op} = 34$ SMEFT degrees of freedom considered in the analysis. To illustrate this point, in table 6 we indicate the sensitivity of each of the LHC processes included in the present analysis along with the degrees of freedom in our fitting basis (for their definition, see table 1). A check mark outside (inside) brackets indicates that a given process constrains the corresponding operator $\mathcal{O}(\Lambda^{-2})$ ($\mathcal{O}(\Lambda^{-4})$). A check mark in square brackets indicates that the operator enters at $\mathcal{O}(\Lambda^{-2})$ but only at NLO.

JHEP04(2019)100

Process	SM	Code	SMEFT	Code
$t\bar{t}$	NNLO QCD	MCFM/SHERPA NLO + NNLO K-factors	NLO QCD	MG5_aMC
single-t (t-ch)	NNLO QCD	MCFM NLO + NNLO K-factors	NLO QCD	MG5_aMC
single-t (s-ch)	NLO QCD	MCFM	NLO QCD	MG5_aMC
tW	NLO QCD	MG5_aMC	NLO QCD	MG5_aMC
tZ	NLO QCD	MG5_aMC	LO QCD + NLO SM K-factors	MG5_aMC
$t\bar{t}W(Z)$	NLO QCD	MG5_aMC	LO QCD + NLO SM K-factors	MG5_aMC
$t\bar{t}h$	NLO QCD	MG5_aMC	LO QCD + NLO SM K-factors	MG5_aMC
$t\bar{t}t\bar{t}$	NLO QCD	MG5_aMC	LO QCD + NLO SM K-factors	MG5_aMC
$t\bar{t}b\bar{b}$	NLO QCD	MG5_aMC	LO QCD + NLO SM K-factors	MG5_aMC

Table 5. Summary of the theoretical calculations used for the description of the LHC top production cross-sections included in the present analysis. We indicate, for both the SM and the SMEFT contributions to the cross-sections, the perturbative accuracy and the codes used to produce the corresponding predictions.

The comparison in table 6 illustrates the importance of a global approach to the SMEFT analysis of top quark production. On the one hand, several operators are constrained by many different processes, and this allows independent and complementary constraints. For instance, the chromomagnetic operator c_{tG} is relevant for the description of all the input processes with the exception of single-top production. On the other hand, other operators are constrained by one or two processes at most, so that information on them can be obtained only by including a wide range of different input observables. For instance, $c_{t\varphi}$, constrained only by $t\bar{t}H$ production; c_{bW}, sensitive only to single-top production at $\mathcal{O}(\Lambda^{-4})$; and the four-heavy-quark operators for which the only available information is from $t\bar{t}b\bar{b}$ and $t\bar{t}t\bar{t}$.

From table 6 we also observe that adding the formally subleading $\mathcal{O}(\Lambda^{-4})$ contributions from the dimension-six operators increases the sensitivity of many different processes. For example, the ObW and Off operators can only be constrained once $\mathcal{O}(\Lambda^{-4})$ terms are included in the fit. This is also true for several of the four-fermion operators, for which additional constraints can be obtained from the $t\bar{t}$, $t\bar{t}V$ and $t\bar{t}H$ production processes once $\mathcal{O}(\Lambda^{-4})$ corrections are taken into account.

Needless to say, it is in principle inconsistent to account only for the $\mathcal{O}(\Lambda^{-4})$ effects arising from the dimension-six operators and not from the dimension-eight operators. However, there are good reasons why it could be worth including them in the analysis. First, including or not the $\mathcal{O}(\Lambda^{-4})$ corrections provides an estimate of whether the SMEFT fit

JHEP04(2019)100

Notation	Sensitivity at $\mathcal{O}(\Lambda^{-2})$ ($\mathcal{O}(\Lambda^{-4})$)								
	$t\bar{t}$	single-top	tW	tZ	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}t\bar{t}$	$t\bar{t}b\bar{b}$
OQQ1								✓	✓
OQQ8								✓	✓
OQt1								✓	✓
OQt8								✓	✓
OQb1									✓
OQb8									✓
Ott1								✓	
Otb1									✓
Otb8									✓
OQtQb1									(✓)
OQtQb8									(✓)
O81qq	✓				✓	✓	✓	✓	✓
O11qq	[✓]				[✓]	[✓]	[✓]	✓	✓
O83qq	✓	[✓]		[✓]	✓	✓	✓	✓	✓
O13qq	[✓]	✓		✓	[✓]	[✓]	[✓]	✓	✓
O8qt	✓				✓	✓	✓	✓	✓
O1qt	[✓]				[✓]	[✓]	[✓]	✓	✓
O8ut	✓					✓	✓	✓	✓
O1ut	[✓]					[✓]	[✓]	✓	✓
O8qu	✓					✓	✓	✓	✓
O1qu	[✓]					[✓]	[✓]	✓	✓
O8dt	✓					✓	✓	✓	✓
O1dt	[✓]					[✓]	[✓]	✓	✓
O8qd	✓					✓	✓	✓	✓
O1qd	[✓]					[✓]	[✓]	✓	✓
OtG	✓		✓		✓	✓	✓	✓	✓
OtW		✓	✓	✓					
ObW		(✓)	(✓)	(✓)					
OtZ				✓		✓			
Off		(✓)	(✓)	(✓)					
Ofq3		✓	✓	✓					
OpQM				✓		✓			
Opt				✓		✓			
Otp							✓		

Table 6. The sensitivity of each of the LHC processes included in the present analysis with each of the $N_{\rm obs} = 34$ dimension-six SMEFT operators that constitute our fitting basis (see table 1 for their definition). A check mark outside (inside) brackets indicates that a given process constrain the corresponding operator $\mathcal{O}(\Lambda^{-2})$ ($\mathcal{O}(\Lambda^{-4})$), and operators in square brackets indicate that an operator contributes at $\mathcal{O}(\Lambda^{-2})$ but only at NLO, see text for details.

JHEP04(2019)100

results are stable upon higher orders in the effective field theory parameter expansion. Second, one might consider scenarios where the dimension-eight operators do not interfere with the SM amplitudes. In this case the only physically relevant $\mathcal{O}(\Lambda^{-4})$ effects are those arising from the dimension-six operators, which we include here.

4 The SMEFiT fitting methodology

In this section, we describe the SMEFiT fitting approach that we adopt here to constrain the SMEFT operators summarised in table 1. First, we explain how the MC replica method can be used to construct the probability density in the space of the fitted SMEFT coefficients $\{c_i\}$. Then, we motivate the choice of the figure of merit used for the minimisation, discuss the propagation of experimental and theoretical uncertainties, and explain how we determine the best-fit parameters in a way that avoids over-fitting. Finally, we describe how the fitting methodology can be validated by means of closure tests, analogously to the PDF case, and apply this strategy to study the robustness of the results and their dependence with a number of fit settings.

4.1 The Monte Carlo replica method

In this work, we adopt the MC replica method to propagate the experimental uncertainties from the input experimental cross-sections to the fitted SMEFT coefficients $\{c_i\}$. The idea underlying this method is to construct a sampling of the probability distribution in the space of the experimental data, which then translates into a sampling of the probability distribution in the space of the SMEFT coefficients by means of the fitting procedure. This strategy can be implemented by generating a large number ($N_{\rm rep}$) of artificial replicas of the original data. The replica generation is based on the available information on the experimental central values, uncertainties, and correlations associated to each of the input data points. It can then be shown that averages, variances, and correlations computed over the sample of $N_{\rm rep}$ MC replicas reproduce the corresponding experimental values.

In practice, the MC replica method works as follows. Given an experimental measurement of a hard-scattering cross-section, denoted by $\mathcal{O}_i^{(\rm exp)}$, with total uncorrelated uncertainty $\sigma_i^{(\rm stat)}$, $N_{\rm sys}$ correlated systematic uncertainties $\sigma_{i,\alpha}^{(\rm sys)}$, and $N_{\rm norm}$ normalisation uncertainties $\sigma_{i,n}^{(\rm norm)}$ (such as those associated to the luminosity), the artificial replicas are generated as

$$\mathcal{O}_i^{(\rm art)(k)} = S_{i,N}^{(k)} \mathcal{O}_i^{(\rm exp)} \left(1 + r_i^{(k)} \sigma_i^{(\rm stat)} + \sum_{\alpha=1}^{N_{\rm sys}} r_{i,\alpha}^{(k)} \sigma_{i,\alpha}^{(\rm sys)} \right), \quad k = 1, \ldots, N_{\rm rep}\,, \tag{4.1}$$

where the index i runs from 1 to $N_{\rm dat}$, the total number of points in a specific dataset, and where the normalisation prefactor is given by

$$S_{i,N}^{(k)} \equiv \prod_{n=1}^{N_{\rm norm}} \left(1 + r_{i,n}^{(k)} \sigma_{i,n}^{(\rm norm)} \right). \tag{4.2}$$

JHEP04(2019)100

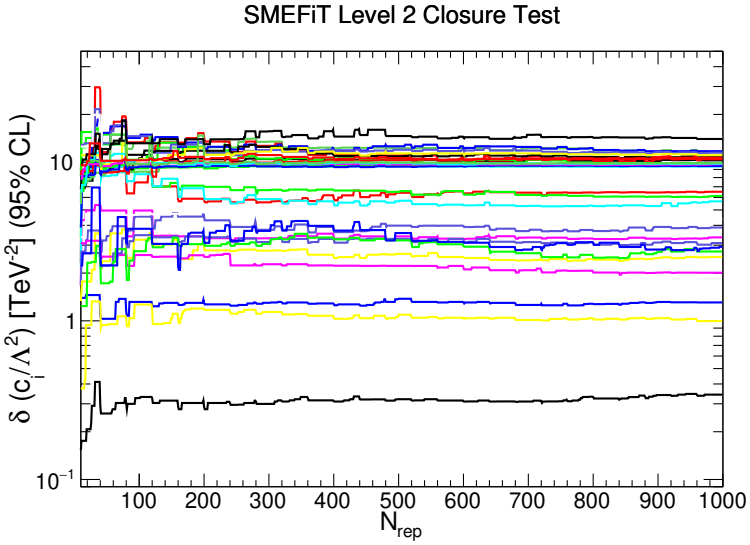

Figure 2. The dependence of the 95% CL bounds on the fit parameters $\delta c_i/\Lambda^2$ with the number of MC replicas $N_{\rm rep}$ used in the fit, as determined in a level 2 closure test. Each of the lines shown in the plot correspond to one of the $N_{\rm rep} = 34$ degrees of freedom defined in table 1.

In eqs. (4.1) and (4.2), $r_i^{(k)}$, $r_{i,\alpha}^{(k)}$, and $r_{i,n}^{(k)}$ are univariate Gaussian random numbers. Correlations between data points induced by systematic uncertainties are accounted for by ensuring that $r_{i,\alpha}^{(k)} = r_{i',\alpha}^{(k)}$. A similar condition is applied for multiplicative normalisation uncertainties if the n-th normalisation uncertainty is common to the entire dataset, i.e. $r_{i,n}^{(k)} = r_{i',n}^{(k)}$.

The MC approach is conceptually different from the commonly adopted Hessian method, based on the expansion of the χ^2 around its best-fit minimum assuming a quadratic behaviour. Nevertheless, under specific conditions, the two methods can be shown to reproduce equivalent results for the determination of the uncertainties in fitted parameters, see e.g. ref. [23] for studies in the PDF context. The main advantage of the MC method is that it does not require any assumption about the underlying probability distribution of the parameters, and in particular it is not restricted to Gaussian distributions. Moreover, it is suited to problems where the parameter space is large and complicated, with a large number of quasi-degenerate minima and flat directions. For these reasons, adopting the MC approach rather than the Hessian method is rather advantageous in the case of SMEFT fits.

An important aspect to address in the MC method is how many replicas $N_{\rm rep}$ need to be generated for each specific application. In order to determine this, we assess the robustness of our results with respect to the number of MC replicas used in the fit. To do so, in figure 2 we show the dependence of the bounds $\delta c_i/\Lambda^2$, determined at the 95% confidence level, on the value of $N_{\rm rep}$ from a level 2 closure test, discussed in detail in section 4.4. Each line corresponds to one of the $N_{\rm op} = 34$ degrees of freedom defined in table 1. From figure 2, we find that for $N_{\rm rep} \lesssim 100$ the fit estimate for the bounds is affected by large fluctuations. These fluctuations are dampened as the number of replicas

JHEP04(2019)100

increases, and for $N_{\rm rep} \gtrsim 500$ the results become independent from $N_{\rm rep}$. In order to ensure that no residual MC fluctuations remain, we will use $N_{\rm rep} = 1000$ as our baseline. We note however that the validity of this conclusion, in general, will depend on the input dataset, and should therefore be reconsidered if this is modified, in particular if the dataset is significantly extended.

4.2 χ^2 definition

For each of the MC replicas generated with eq. (4.1), we perform a fit to the $N_{\rm op} = 34$ degrees of freedom c_l/Λ^2, $l = 1, \ldots, N_{\rm op}$, associated to the dimension-6 SMEFT operators defined in section 2. This process results in a set of $\{c_l^{(k)}\}$ best-fit values for each replica, from which estimators such as expectation values, variances and correlations can be readily evaluated.

For each MC replica, the corresponding best-fit values are determined from the minimisation of a figure of merit, the error function, defined as

$$E(\{c_l^{(k)}\}) \equiv \frac{1}{N_{\rm dat}} \sum_{i,j=1}^{N_{\rm dat}} \left(\mathcal{O}_i^{(\rm th)}\left(\{c_l^{(k)}\}\right) - \mathcal{O}_i^{(\rm art)(k)}\right)(\mathrm{cov}^{-1})_{ij}\left(\mathcal{O}_j^{(\rm th)}\left(\{c_l^{(k)}\}\right) - \mathcal{O}_j^{(\rm art)(k)}\right) , \tag{4.3}$$

where $N_{\rm dat}$ is the number of data points used in the fit, and $\mathcal{O}_i^{(\rm th)}$ is the theoretical prediction for the i-th cross-section evaluated using the $\{c_l^{(k)}\}$ values for the SMEFT degrees of freedom. Note that in eq. (4.3) the theory predictions are compared to the MC replicas, rather than to the original experimental central values. Once the best-fit parameters have been determined for all the $N_{\rm rep}$ replicas, the overall fit quality can be quantified by means of the χ^2

$$\chi^2 \equiv \frac{1}{N_{\rm dat}} \sum_{i,j=1}^{N_{\rm dat}} \left(\mathcal{O}_i^{(\rm th)}(\{\langle c_l\rangle\}) - \mathcal{O}_i^{(\rm exp)}\right)(\mathrm{cov}^{-1})_{ij}\left(\mathcal{O}_j^{(\rm th)}(\{\langle c_l\rangle\}) - \mathcal{O}_j^{(\rm exp)}\right) , \tag{4.4}$$

where now the theoretical predictions, computed using the expectation value (the mean) for the degree of freedom c_l, are compared to the central experimental data. This is evaluated as the average over the resulting MC best-fit sample $\{c_l^{(k)}\}$

$$\langle c_l\rangle \equiv \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} c_l^{(k)} . \tag{4.5}$$

Both the error function, eq. (4.3), and the χ^2, eq. (4.4), are expressed in terms of the total covariance matrix, cov_{ij}, which should contain all the relevant sources of experimental and theoretical uncertainties. Assuming that theoretical uncertainties follow an underlying Gaussian distribution, and that they are uncorrelated to the experimental uncertainties, it can be shown [158] that the total covariance matrix can be expressed as

$$\mathrm{cov}_{ij} = \mathrm{cov}_{ij}^{(\rm exp)} + \mathrm{cov}_{ij}^{(\rm th)} , \tag{4.6}$$

that is, as the sum of the experimental and theoretical covariance matrices.

JHEP04(2019)100

Concerning the experimental covariance matrix, we use the so-called 't_0' definition [159]

$$(\mathrm{cov}_{t_0})_{ij}^{(\mathrm{exp})} \equiv \left(\sigma_i^{(\mathrm{stat})}\right)^2 \delta_{ij} + \left(\sum_{\alpha=1}^{N_{\mathrm{sys}}} \sigma_{i,\alpha}^{(\mathrm{sys})} \sigma_{j,\alpha}^{(\mathrm{sys})} \mathcal{O}_i^{(\mathrm{exp})} \mathcal{O}_j^{(\mathrm{exp})} + \sum_{\beta=1}^{N_{\mathrm{norm}}} \sigma_{i,\beta}^{(\mathrm{norm})} \sigma_{j,\beta}^{(\mathrm{norm})} \mathcal{O}_i^{(\mathrm{th},0)} \mathcal{O}_j^{(\mathrm{th},0)} \right), \tag{4.7}$$

where one treats the additive ('sys') relative experimental systematic errors separately from the multiplicative ('norm') ones. In the additive case, one uses the central value of the experimental measurement, $\mathcal{O}_i^{(\mathrm{exp})}$. In the multiplicative case, one uses instead a fixed set of theoretical predictions, $\{\mathcal{O}_i^{(\mathrm{th},0)}\}$. These theoretical predictions are typically obtained from a previous fit; the fit is then iterated until consistency is reached. The use of the t_0 covariance matrix defined in eq. (4.7) avoids the bias associated to multiplicative uncertainties, which would lead to a systematic undershooting of the best-fit values as compared to their true values [160].

As mentioned in section 3, we construct the experimental covariance matrix, eq. (4.7), from all available sources of statistical and systematic uncertainties for a given dataset. Information on the bin-by-bin correlations of systematic uncertainties is available only for a subset of the data listed in tables 2–3, specifically, for all the 8 TeV top-quark pair differential distributions in table 2 and for the corresponding CMS distributions at 13 TeV from ref. [98]. For all the other measurements, we add all uncertainties in quadrature; our analysis can be easily updated should more correlations become available.

In addition to the experimental uncertainties, there are at least two main classes of theoretical uncertainties that are in principle relevant for the present fits: (i) uncertainties associated to missing higher orders (MHOs) in the perturbative calculation, and (ii) PDF uncertainties. The impact of the former is not expected to affect this analysis significantly, because we perform the SM calculation at the highest available perturbative order. In particular, we take into account NNLO QCD corrections for the two families of processes that are more precisely known experimentally, namely the absolute differential distributions in inclusive $t\bar{t}$ and single top (t-channel) production. Furthermore, as discussed in section 2.4, for most of the SMEFT contributions, the NLO QCD calculation is used.

The inclusion of PDF uncertainties, instead, is more important. In this work, we use as input to all our theory calculations the NNPDF3.1 NNLO no-top PDF set [36], which differs from the NNPDF3.1 baseline set only for the exclusion of the top-quark pair production data from the dataset. As explained in section 3, this is necessary to avoid double-counting in the fit. However, this implies that the SM calculation of top quark pair production could be affected by sizeable PDF uncertainties, especially in the tails of the differential distributions, which are not constrained by alternative gluon-sensitive processes in the fit such as transverse momentum Z-boson [161], jet [162], and direct photon production [163]. Therefore, not accounting for PDF uncertainties may bias the results of the fit.

JHEP04(2019)100

With this motivation, we construct the theoretical covariance matrix from the contributions of the PDF uncertainty as

$$\text{cov}_{ij}^{(\text{th})} = \left\langle \mathcal{O}_i^{(\text{th})(\text{r})} \mathcal{O}_j^{(\text{th})(\text{r})} \right\rangle_{\text{rep}} - \left\langle \mathcal{O}_i^{(\text{th})(\text{r})} \right\rangle_{\text{rep}} \left\langle \mathcal{O}_j^{(\text{th})(\text{r})} \right\rangle_{\text{rep}}, \tag{4.8}$$

where the theoretical predictions $\mathcal{O}_i^{(\text{th})(\text{r})}$ are computed using the SM theory and the r-th replica from the NNPDF3.1NNLO no-top PDF set, and averages $\langle\cdot\rangle_{\text{rep}}$ are performed over the $\widetilde{N}_{\text{rep}} = 100$ replicas of this PDF set. Note that replicas in the PDF set are not directly related to replicas in the SMEFT set, since the two sets represent different probability distributions.

In general, the theoretical covariance matrix, eq. (4.8), induces correlations between all the datasets included in the fit. However we account for them only within a given dataset, in the same way as for experimental measurements. If the PDF-induced correlations between data points i and j are neglected, eq. (4.8) reduces to

$$\text{cov}_{ii}^{(\text{th})} = \left\langle \left(\mathcal{O}_i^{(\text{th})(\text{k})} \right)^2 \right\rangle_{\text{rep}} - \left\langle \mathcal{O}_i^{(\text{th})(\text{k})} \right\rangle_{\text{rep}}^2, \tag{4.9}$$

and vanishes for $i \neq j$. This corresponds to adding the PDF error in quadrature to the experimental uncertainties.

For consistency, PDF uncertainties should be included in the fit not only via the covariance matrix in eqs. (4.3)–(4.4), but also in the MC replica generation. That is, the generation of the data replicas according to eq. (4.1) includes an additional source of fluctuation determined from the theoretical covariance matrix, eq. (4.8). Note that, for the k-th data replica, the theory predictions $\mathcal{O}_i^{(\text{th})}(\{c_l^{(k)}\})$ are evaluated using a different PDF replica from the NNPDF3.1NNLO no-top set. Since in general the number of data replicas, $N_{\text{rep}} = 1000$, is much larger than the number of PDF replicas, $\widetilde{N}_{\text{rep}} = 100$, the latter are selected at random with repetition for each data replica.

4.3 Minimisation and stopping

In the case of current SMEFT fits, the minimisation of the error function, E, eq. (4.3), may be achieved by exploiting gradient descent methods, which rely on variations of E. This is because the relationship between the theory cross-sections and the fitted parameters is at most quadratic, see eq. (2.2). Taking this into account, the optimiser that we use here to determine the best-fit values of the degrees of freedom $\{c_i\}$ is the sequential least squares programming algorithm `SLSQP` [164] available in the `SciPy` package. It belongs to the family of sequential quadratic programming methods, which are based on solving a sequence of optimisation subproblems, where each of them optimises a quadratic model. An advantage of using `SLSQP` is that it allows one to provide the optimiser with any combination of constraints on the coefficients, including existing bounds, a feature that might become useful for future studies.

Since the dimensionality of this parameter space is not that different from the total number of input cross-sections ($N_{\text{dat}} = 103$ points), one needs to avoid over-fitting, i.e. fitting the statistical fluctuations of the experimental data rather than the underlying physical

JHEP04(2019)100

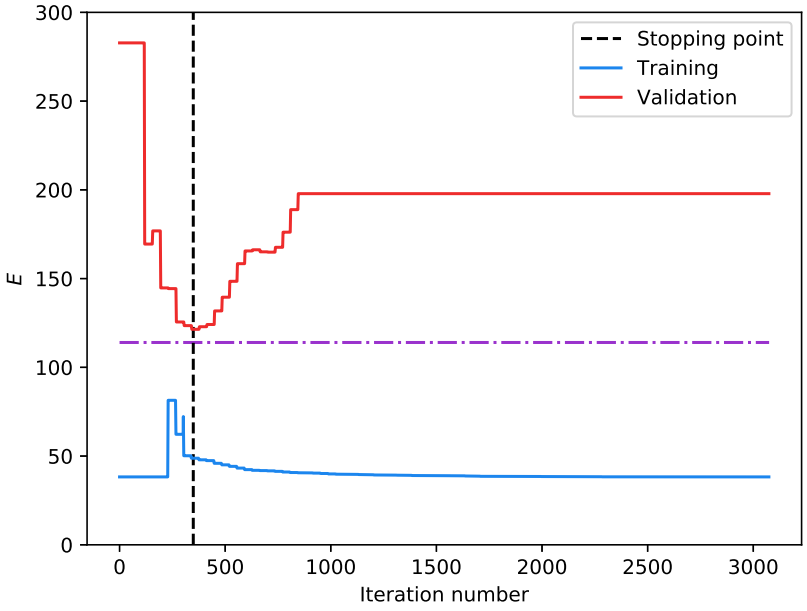

Figure 3. The values of the training and validation error functions, $E_{\rm tr}$ and $E_{\rm val}$ respectively, as a function of the number of iterations of the minimisation algorithm for a representative data replica. The optimal stopping point, as determined by the cross-validation algorithm, is indicated by the dashed vertical line. The horizontal dot-dashed line indicates the value $E/N_{\rm dat} = 2$, see text.

law. Such an effect is particularly dangerous in a situation like the current one, where there are a large number of flat directions with several parameters strongly (anti)-correlated.

To prevent the minimiser from over-fitting the data, we use (MC) cross-validation. For each replica, the data is randomly split with equal probability into two disjoint sets, known as the training and validation sets. Only the data points in the training set are then used to compute the figure of merit being minimised, eq. (4.3), while the data points in the validation set are monitored alongside the fit. The random assignment of the data points to the training or validation sets is different for each MC replica.

The optimal stopping point of the fit is reached when the figure of merit evaluated on the validation set, $E_{\rm val}^{(k)}$, starts to increase. This is not equivalent to the absolute minimum of the error function evaluated on the training set, $E_{\rm tr}^{(k)}$. To illustrate how cross-validation works, in figure 3 we show the values of the training and validation error functions, $E_{\rm tr}$ and $E_{\rm val}$ respectively, as a function of the number of iterations of the minimisation algorithm for a representative data replica. The optimal stopping point is indicated by the dashed vertical line. For reference, the horizontal dot-dashed line indicates the value $E/N_{\rm dat} = 2$, corresponding to the expectation value of the error function close to the optimal fit. Note that $E \sim 1$ would indicate over-fitting as well, since we have two independent sources of fluctuations in the artificial data: the original data fluctuations, and the fluctuations induced by the MC replicas.

From this comparison, we can observe the expected behaviour for both the training and validation sets; namely that the validation χ^2, once it reaches its lowest value, increases rapidly, while the training χ^2 continues to decrease. We find $E_{\rm tr} \ll E_{\rm val}$ as the number of

JHEP04(2019)100

iterations increase, which is an indication that the optimisation algorithm is over-fitting. It is therefore clear that, without adopting cross-validation, the absolute minimum found by the optimisation algorithm would not correspond to the true underlying law, but rather to fitting statistical noise. We will quantify the importance of cross-validation in the SMEFT fit results in section 4.5, where we will show that without it one obtains unreliable results, including spurious deviations from the SM predictions.

A peculiar behaviour in figure 3 is that the cost function E remains flat for the initial part of the minimisation. This feature arises due to the fact that for some replicas the (randomly chosen) initial starting point of the fit may either turn out to be close enough to the Standard Model values, or to induce compensations between operators that effectively lead to a cost function E which is similar to the SM one. For those replicas for which the initial value of E is higher than the SM result, we observe the expected behaviour where the training cost function decreases rapidly at low iteration numbers.

4.4 Closure test validation

A reliable fitting framework should be able to fit a wide range of different datasets without tuning the methodology and without biasing the results. Validating a new methodology can be complicated by issues such as potential inconsistencies (internal or external) in the experimental data, or by limitations in the theoretical calculations. To validate the fitting methodology used in this SMEFT analysis, we carry out a series of closure tests, based on pseudo-data generated with a known underlying physical law, see [33] for more details.

The basic idea underlying a closure test is to test the SMEFT fitting procedure by performing fits where the "correct" result is known, i.e. by fitting pseudo-data generated from a fixed reference set of values for the SMEFT degrees of freedom, $\{c_i^{(\mathrm{ref})}\}$. Closure tests allow one to check that the fitting methodology can reproduce the underlying law, which is known by construction. The SM and SMEFT theory calculations can be assumed to be exact, since we use the same theory settings to generate and fit the pseudo-data. As a consequence, the theoretical uncertainties associated to MHOs and PDFs do not enter closure tests, where only methodological and experimental uncertainties are checked. In the case of SMEFT fits, we can perform a closure test assuming that the underlying truth is the SM, i.e. $\{c_i^{(\mathrm{ref})} = 0\}$, or any BSM scenario, i.e. $\{c_i^{(\mathrm{ref})} \neq 0\}$. This allows for the validation of potential BSM anomalies identified in the SMEFT fit to the actual experimental data.

In the following, we consider three levels of closure tests according to the type of pseudo-data that is used as input to the fit.

- In a level zero (L0) closure test the pseudo-data coincides with the true underlying law, without any additional fluctuations. Then N_{rep} fits are performed to exactly the same pseudo-data, with the only difference being the random initial conditions in each case. For instance, if the pseudo-data is generated with the SM hypothesis $\{c_i^{(\mathrm{ref})} = 0\}$, then the same values should be reproduced at the fit level within uncertainties. For a L0 closure test, the training/validation partition is not necessary, since the information contained in both sets would be identical.

JHEP04(2019)100

In a L0 closure test, one expects the error function E to tend to zero for a large enough number of iterations. Therefore, direct evidence that a L0 test is successful is to show how the error function decreases with the number of iterations. A L0 closure test therefore allows one to check that the minimiser is efficient enough to properly explore the entire parameter space.

- In a level one (L1) closure test, one adds noise on top of the pseudo-data. Two types of noise may be added; in a L1a closure test we generate MC replicas of the pseudo-data generated in a L0 closure test in the same way as in a real fit to data. Alternatively, in a L1b closure test one adds stochastic noise directly to the pseudo-data, in order to replicate experimental uncertainties included in the fits to data. In this work we adopt a L1a-type closure test, in contrast to NNPDF, where L1b-type closure tests are used as a default. We note that adopting a L1a-type test over L1b simply means that a different type of uncertainty is being probed at L1 — we discuss the various types of uncertainty in section 4.5 where we characterise the types of fit uncertainties.

  In comparison to a L0 closure test, a L1 closure test propagates the experimental uncertainties into the fitted coefficients, and can therefore be used to demonstrate that the quoted uncertainties in the fit parameters admit a robust statistical interpretation. One expects $E \sim 1$ for a successful closure test.

- In a level two (L2) closure test, one adds the aforementioned stochastic noise on top of the MC replicas included in the L1 closure test. This statistical noise is generated according to the experimental covariance matrix of the real data. A L2 closure test is therefore equivalent to a fit to the real data, the only difference being that data and theory are perfectly consistent by construction.

L0 closure tests. First of all, we want to demonstrate that the optimiser is efficient enough to explore the full 34-dimensional parameter space. With this motivation, pseudo-data corresponding to the SM has been generated for all the cross-sections described in tables 2–4 and fitted without introducing any additional noise. As mentioned above, here all data is fitted since the training/validation separation is not required. In figure 4 we show the error function E for L0 closure tests based on the SM scenario as a function of the number of iterations in the minimiser for three replicas with different initial boundary conditions. We see how the error function decreases with the number of iterations, approaching the limit $E \to 0$ which corresponds to the case where the fit results reproduce the reference values $\{c_i = 0\}$.

In figure 5 we show the results of the L0 closure tests; in the left plot we show the fit residuals for the $N_{\rm op} = 34$ degrees of freedom included in the fit. They are defined as

$$r_i \equiv \frac{\left(\langle c_i \rangle - c_i^{(\rm ref)}\right)}{\delta c_i}\,, \tag{4.10}$$

where $\langle c_i \rangle$ and δc_i indicate the expectation value, eq. (4.5), and the 95% CL.

JHEP04(2019)100

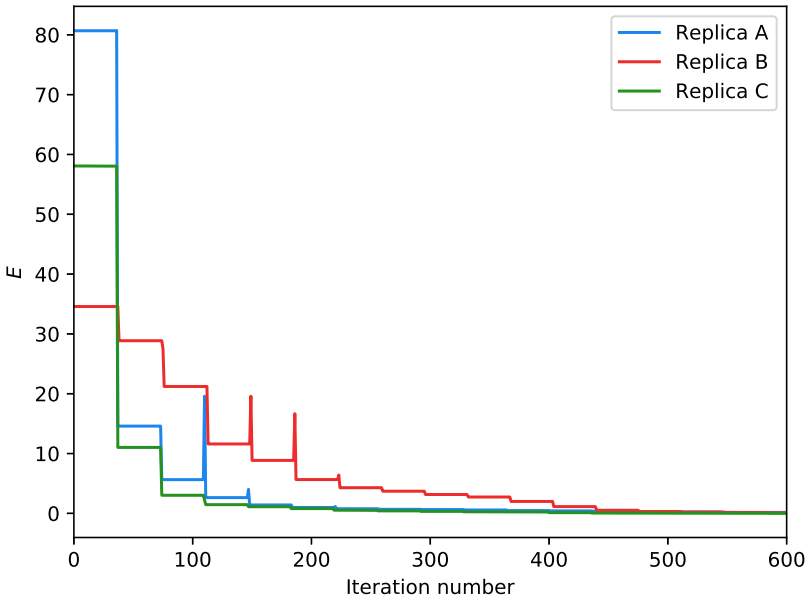

Figure 4. The error function E for L0 closure tests based on the SM scenario as a function of the number of iterations. We show the results for three fits to the same pseudo-data but with different initial boundary conditions.

As before, $c_i^{(\mathrm{ref})}$ represent the reference values of the SMEFT degrees of freedom used to generate the pseudo-data, which here are set to zero. We find that the residuals are all very close to zero, i.e. the optimiser has managed to identify with good accuracy the true underlying values of the fit parameters.

In the same figure, we also show the corresponding values of the 95% CL on the fit parameters δc_i. The units of the δc_i are TeV^{-2}, and as in the rest of this work for reference we are assuming that $\Lambda = 1\,\mathrm{TeV}$. While the bounds δc_i span up to two orders of magnitude, in all cases they will be smaller or comparable to those obtained from the corresponding Level 2 closure tests, to be discussed below. Indeed, as will be shown in figure 11, the L0 bounds can be an order of magnitude smaller (or even more) for specific degrees of freedom.

L2 closure tests. In this case we compare the results based on two different hypotheses for the underlying theory; one in which all degrees of freedom are set to their SM values; and another in which all degrees of freedom are set to their SM values, except $c_{tu}^8/\Lambda^2 = 20\,\mathrm{TeV}^{-2}$. The goal of this exercise is to verify if the fit can successfully identify a BSM deviation once it is built into the pseudo-data. The results of the L2 closure test to the SM is shown in figure 6; we find that the fit manages to reproduce the underlying law, since the residuals r_i are very close to zero indicating that the fit results $\langle c_i \rangle$ agree with the corresponding references values $c_i^{(\mathrm{ref})}$.

Another important output of the L2 closure tests is the size of the bounds on the SMEFT degrees of freedom reported in figure 6 (right). These bounds reflect the constraints on the fit parameters that can be expected based on the input experimental dataset in the case of perfect consistency between data and theory. They provide a baseline to later

JHEP04(2019)100

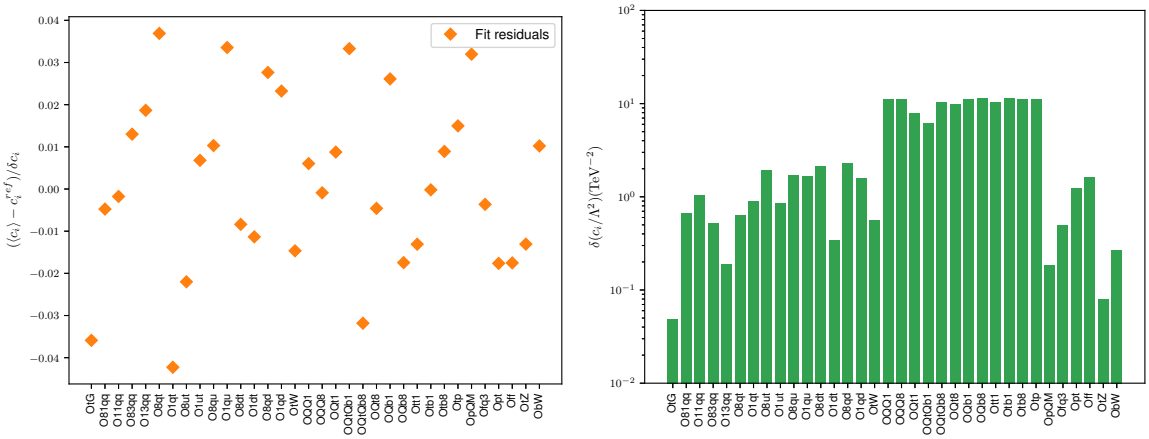

Figure 5. Left: the fit residuals, defined as in eq. (4.10) for the 34 SMEFT degrees of freedom included in the fit for the L0 closure test. Right: the corresponding values of the 95% CL for the fit parameters δc_i. Note that since in general the bounds on each degree of freedom will be rather different, this comparison is performed on a log scale.

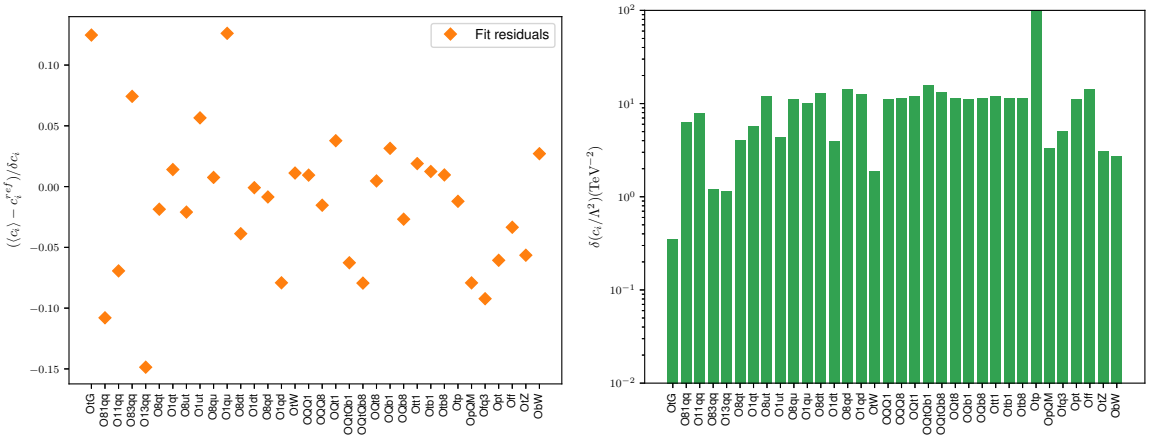

Figure 6. Same as in figure 5, but now for the L2 closure tests.

compare the corresponding results at the level of the fits to the real data. We see from this comparison that some degrees of freedom will be constrained rather better than others: for instance one expects the bounds on `OtG` to be in the range of $\delta c_i \simeq 0.1\,\mathrm{TeV}^{-2}$, while the bounds on `Otp` to be in the range $\delta c_i \simeq 100\,\mathrm{TeV}^{-2}$.

In figure 7 we show the values of the SMEFT degrees of freedom and their uncertainties, $\langle c_i \rangle \pm \delta c_i$, for the L2 closure test in the BSM scenario where one has set $c_{tu}^8/\Lambda^2 = 20\,\mathrm{TeV}^{-2}$, which is approximately twice as large as the 95% CL found by the L2 closure test to the SM. This allows us to ensure the starting point of the fit is BSM for this operator, but is not so far away from the SM as to make the closure test redundant. We firstly observe that the closure test does indeed find a best-fit value for $c_{tu}^8/\Lambda^2 \approx 13\,\mathrm{TeV}^{-2}$, which is outside the error for this operator reported in the SM closure test. If one computes the fit residuals, eq. (4.10) we find that the central value lies outside the 95% CL, which roughly corresponds to 2σ. Therefore at the fit level we would expect to find at least a 2σ deviation

JHEP04(2019)100

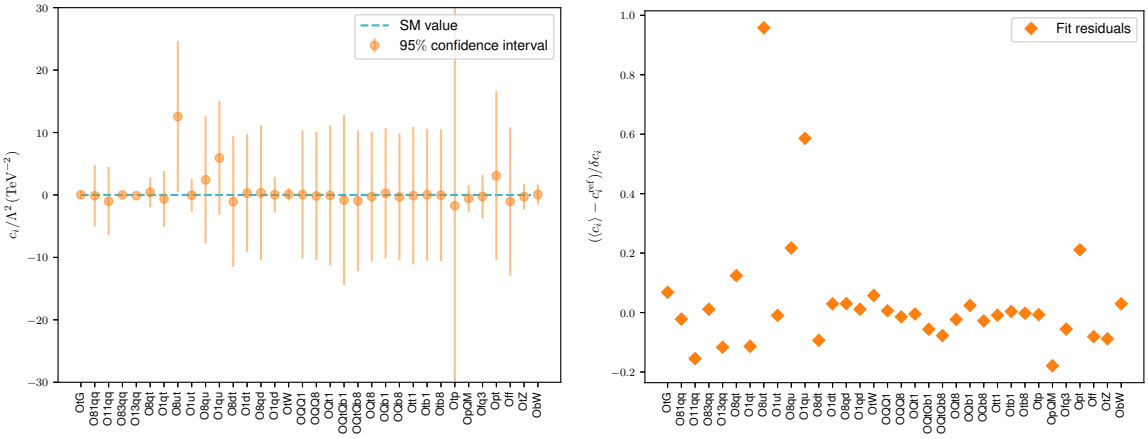

Figure 7. Left: the values of the SMEFT degrees of freedom and their uncertainties, $\langle c_i \rangle \pm \delta c_i$, for the L2 closure test in the BSM scenario where one has set $c_{tu}^8/\Lambda^2 = 20\,\text{TeV}^{-2}$. Right: the corresponding fit residuals.

from the SM, and larger if the value of the coefficient is much larger than the size of the error associated to it.

It is however important to emphasise that the bounds reported in figures 6, 7 need to be taken with a grain of salt, since some degrees of freedom are highly (anti-)correlated. To quantify this, in figure 8 we show the values of the correlation coefficient between the different degrees of freedom c_i for the L2 closure test with SM reference values. The correlation coefficient between two of the degrees of freedom in the fit c_i and c_j is computed using the standard MC expression, namely

$$\rho(c_i, c_j) = \frac{\frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} c_i^{(k)} c_j^{(k)} - \langle c_i \rangle \langle c_j \rangle}{\delta c_i \delta c_j} \,. \tag{4.11}$$

From this comparison we see that some degrees of freedom are very correlated, for example the chromomagnetic operator `OtG` is highly correlated with the two-heavy-two-light operators `O81qq` and `O11qq`.

One might argue that the residuals reported in figure 6 (left) for the Level 2 closure test appear to be unnaturally small, even taking into account that the bounds δc_i are computed as 95% CL ranges. Indeed, if the $N_{\text{op}} = 34$ degrees of freedom that we are considering here were truly independent random variables, then one would expect to find at least one or two operators with residuals $|r_i| \gtrsim 1$. It is possible to demonstrate that the limited spread of the residuals in figure 6 is a direct consequence of both the correlations between the operators and of the redundancies in the SMEFT parameter space, given the input dataset used in the analysis. To show this, in figure 9 we report the same residuals as in figure 6 but now with the results of $N_{\text{op}} = 34$ independent closure tests where each operator has been constrained separately, setting the contribution from all others to zero. As one can see, fitting one operator at a time results in a greater spread of the residuals, with several instances in which $|r_i| \gtrsim 1$. In other words, the reduced spread of the residuals

JHEP04(2019)100

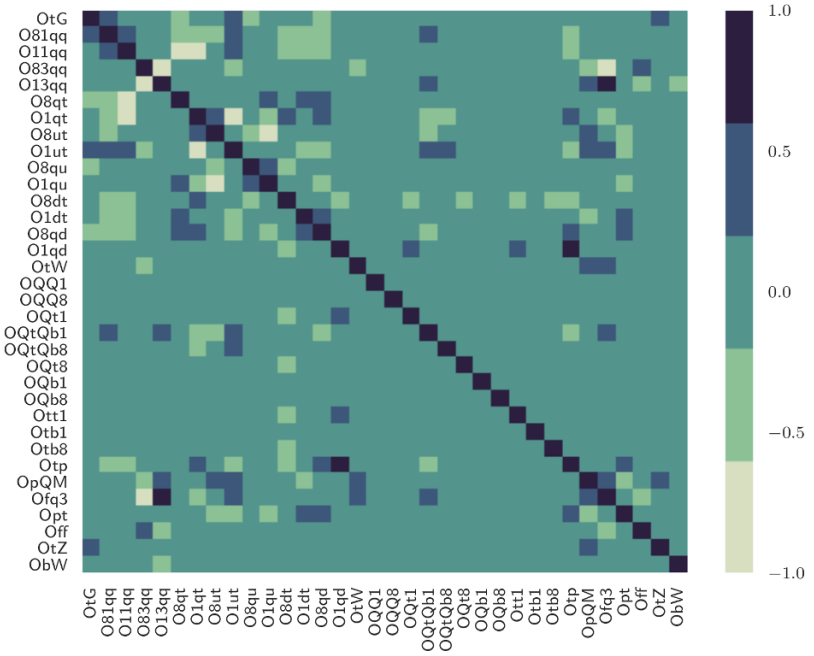

Figure 8. The values of the correlation coefficient between the different degrees of freedom c_i, eq. (4.11), for the L2 closure test with SM reference values.

in the global fit is a genuine effect, arising from the correlations and degeneracies in the explored SMEFT parameter space, rather than an artifact of the fit.

Note that in figure 9 the residuals for the four-heavy-quark operators all take a similar value. The reason is that all these operators are just constrained by the same two data-points, so in individual fit they all result in the same residual. One can check that both $\langle c_i \rangle$ and δ_i are different operator by operator, as expected since in each case the size of the SMEFT corrections is different, but that at the residual level one ends up with the same result.

4.5 Methodological variations

We now turn to study the robustness of the baseline results with respect to a number of variations in the fitting methodology. In particular: (i) the impact of cross-validation; (ii) the effects of experimental uncertainties in determining the bounds on the SMEFT degrees of freedom; and (iii) the role of $\mathcal{O}\left(\Lambda^{-2}\right)$ corrections on these same bounds. We will always assume the SM; as we have shown above, closure tests will likewise work in the case of BSM scenarios.

Cross-validation. As discussed in section 4.3, it is important to ensure that over-fitting is avoided, and, to do so, we adopt cross-validation. To quantify the role that cross-validation plays on fit results, we perform two L2 closure tests, with the only difference that cross-validation is absent in one of them.

In figure 10 we compare the fit residuals and the 95% CL of the fit parameters obtained from the two closure tests. When cross-validation is absent, the central values of the fitted

JHEP04(2019)100

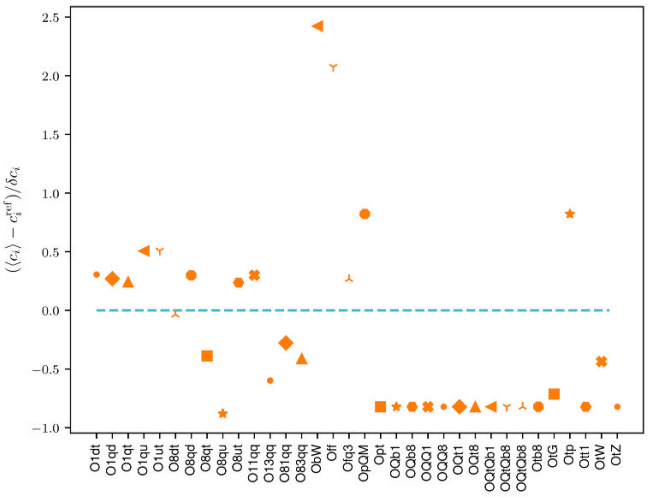

Figure 9. Same as figure 6 (left) for the closure tests in which each of the $N_{\rm op} = 34$ operators has been constrained separately from all the others. The dashed blue line indicates the Standard Model value $\{c_i^{\rm ref}\} = 0$, which is the true underlying law in these closure tests.

degrees of freedom c_i fluctuate around the true result (the SM in this case) rather more than when cross-validation is used. This is a consequence of the fact that the fit without cross-validation has overfitted the experimental data, and therefore the fluctuations around the true result have been enhanced. For example, $r_{\rm bW} \simeq 2.5$ without cross-validation, while it should be $r_{\rm bW} \simeq 0$ as we can see from the left panel. Moreover, from the right panel of figure 10 we see that the bounds obtained by the fit are usually stronger when cross-validation is switched off. However, in this case, they are methodologically biased, and one would incorrectly claim to have derived more stringent limits than the truth. These comparisons highlight that reliable results in a global SMEFT analysis can be obtained only if overfitting is avoided. Otherwise, deviations between experimental data and theory calculations, and/or stringent bounds on the fitted degrees of freedom can be misinterpreted as a sign of new physics, while they are instead a sign of methodological bias.

Characterising fit uncertainties. As explained in section 4.4, L2 closure tests differ from L1 closure tests for the introduction of an additional set of fluctuations. Comparing closure tests at different levels allows one to identify the different components that build up the total uncertainty on the fit parameters δc_i, for a more-in depth discussion applied to PDFs, see ref. [33]. To begin with, L0 closure test results might have *interpolation* and *extrapolation* uncertainties: even if the fit to the data points is perfect, there will be non-zero uncertainties in-between and outside the data region. In the SMEFT case, however these uncertainties vanish in L0 closure tests, since the associated parameter space is discretised over the $N_{\rm op} = 34$ independent degrees of freedom, and additional directions are never explored.

The comparison between the values of δc_i in L1 and L2 closure tests is more subtle. In the L1 case, the data uncertainty is propagated into the fit, see eq. (4.1). Therefore, the component of δc_i, that L1 closure tests identify, is associated to the finite precision of the

JHEP04(2019)100

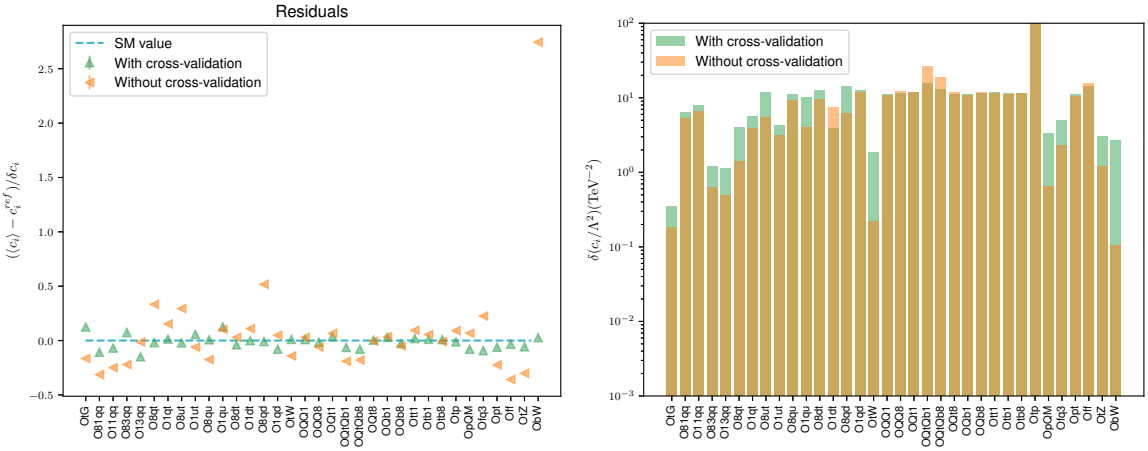

Figure 10. Same as figure 6 comparing the L2 closure tests with and without cross-validation.

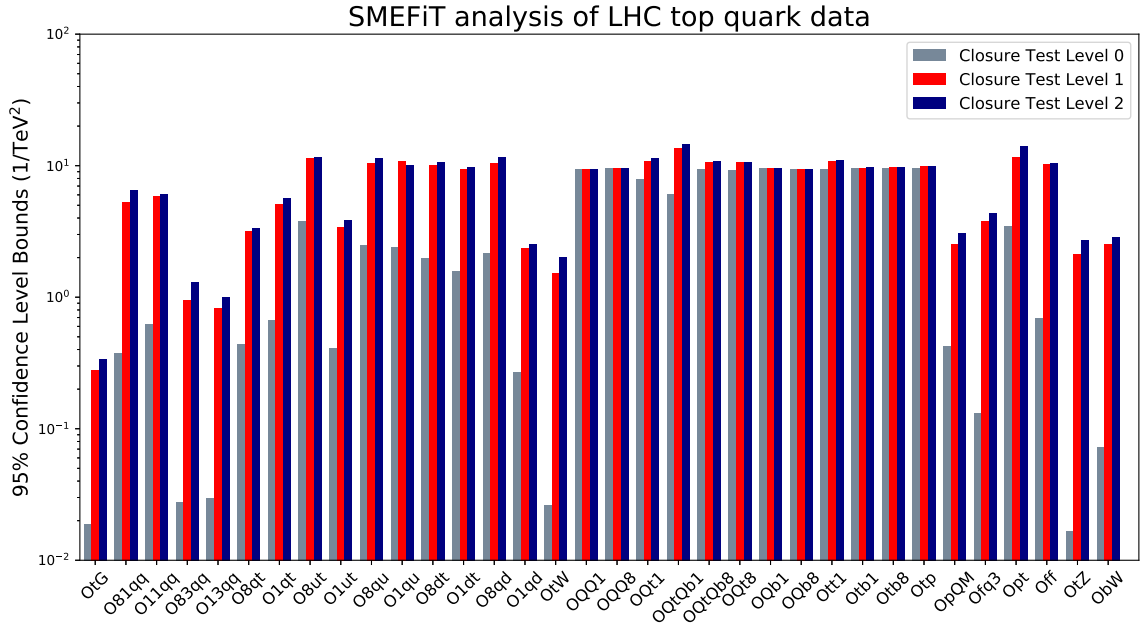

Figure 11. The symmetrised 95% CL bounds for the fit parameters, δc_i, obtained in the closure tests at levels 0, 1a and 2.

input experimental measurements, and hence we call this the *experimental* component of the uncertainty. At L2, we additionally account for the fact that there are infinite different sets of $\{c_i\}$ that optimise the error function equally well. The spread among these solutions represent the irreducible *redundant* component of the uncertainty.

To illustrate the relative weight of these two components on the overall size of δc_i, in figure 11 we show the bounds that are obtained in L0, L1 and L2 closure tests, leaving everything else unchanged. We find that there is a significant increase in the size of δc_i when going from L0 to L1, but then there is only a very slight increase when going from L1 to L2.

JHEP04(2019)100

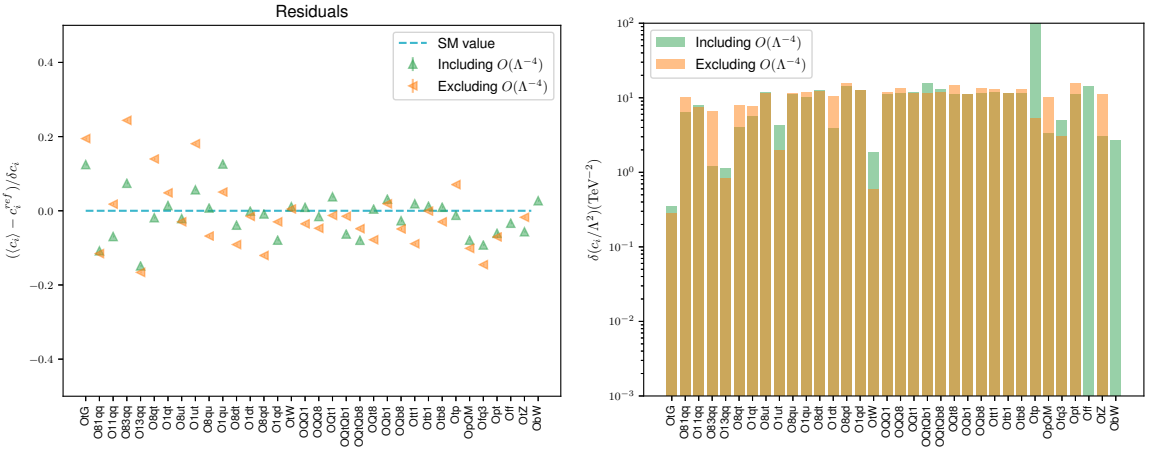

Figure 12. Same as figure 6, now comparing the results of two L2 closure test fits with and without $\mathcal{O}(\Lambda^{-4})$ SMEFT corrections. Our baseline results include these corrections.

The role of $\mathcal{O}(\Lambda^{-4})$ corrections. Closure tests can also be used to assess the dependence of the fit results upon variations of the details of the theory calculations. Specifically, we are interested in the role played by $\mathcal{O}(\Lambda^{-4})$ corrections in the determination of the bounds on the fitted degrees of freedom. As highlighted in table 6, including $\mathcal{O}(\Lambda^{-4})$ terms in the theoretical model modifies rather significantly the parameter space, by opening up new directions and by enhancing the sensitivity to those directions already covered by $\mathcal{O}(\Lambda^{-2})$ terms. Therefore, despite the fact that pseudo-data are generated according to a given theory in a closure test, including or not $\mathcal{O}(\Lambda^{-4})$ corrections implies that the corresponding results should in general be different.

In figure 12 we show the comparison of the residuals r_i (left panel) and of the bounds δc_i (right panel) for L2 closure tests between two fits that differ only for the inclusion (or not) of the $\mathcal{O}(\Lambda^{-4})$ terms. Two degrees of freedom, `Off` and `ObW`, are not constrained in the fit without $\mathcal{O}(\Lambda^{-4})$ terms, and are therefore set to zero. From this comparison, we see that the bounds on the coefficients δc_i generally improve when $\mathcal{O}(\Lambda^{-4})$ corrections are included in the theoretical calculation. For example, the bound on `OtZ` decreases from $\delta c_{tZ} \simeq 6\,\mathrm{TeV}^{-2}$ to $\delta c_{tZ} \simeq 2\,\mathrm{TeV}^{-2}$. The slight worsening observed for the bounds on some few operators when only linear terms are included is consistent with statistical fluctuations, and is therefore not significant. In any case, the fit results are qualitatively similar irrespective of the inclusion of $\mathcal{O}(\Lambda^{-4})$ corrections. Note that some of the degrees of freedom are highly correlated, therefore the interpretation of the results at the individual bound level should be taken with care.

5 The top quark sector of the SMEFT at NLO

In this section, we present the main results of this work, namely we derive the constraints on the $N_{\rm op} = 34$ SMEFT dimension-6 degrees of freedom relevant for the interpretation of top quark production measurements at the LHC. We first discuss the fit quality and the agreement between experimental data and theoretical predictions for individual processes.

JHEP04(2019)100

We then present the best-fit values, the 95% confidence level intervals and the correlations for these degrees of freedom, and we compare our results with other related analyses in the literature. We also study the impact that both NLO QCD perturbative corrections and quadratic $\mathcal{O}\left(\Lambda^{-4}\right)$ terms have on the results. Finally, we assess the dependence of the fit results on the choice of input dataset, and quantify the dependence of the derived bounds on the high-energy limit of the cross-sections included in the fit.

5.1 Fit quality and comparison with data

We will first assess the quality of the fit at the level of both the total dataset and of individual measurements, and then compare the fit results with the input experimental cross-sections. In the following, as discussed in section 2, our baseline fit is based on $N_{\rm rep} = 1000$ MC replicas and includes both NLO QCD corrections for the SMEFT contributions and the quadratic $\mathcal{O}\left(\Lambda^{-4}\right)$ higher order terms.

In table 7 we indicate the values of the χ^2 per datapoint for each of the datasets included in the fit. In each case, we indicate the values of $\chi^2/n_{\rm dat}$ first when the theory calculations include only the SM contributions (second column) and then once they account for the SMEFT corrections after the fit (third column). In the last column, we indicate the number of data points $n_{\rm dat}$ for each dataset. The datasets are classified into three groups following the structure of tables 2–4: inclusive $t\bar{t}$, $t\bar{t}$ in association with V, H, or heavy quarks, and single top production. In the case of datasets consisting of multiple differential distributions, we indicate the one that has been included in this analysis.

From the values in table 7 we find that the overall fit quality to the $n_{\rm dat} = 103$ data points included in the fit is satisfactory, with of $\chi^2/n_{\rm dat} = 1.06$ (1.11) after (before) the fit. We find therefore a slight improvement in the overall fit quality once the dimension-6 SMEFT corrections are taken into account. Note however that this improvement is not inconsistent with statistical fluctuations, since for 103 points one expects $\Delta\left(\chi^2/n_{\rm dat}\right) \simeq 0.1$.

For most of the individual datasets, the SM description of the input measurements is already good to begin with. In several cases, the χ^2 decreases once the SMEFT corrections are accounted for. For instance, the ATLAS $m_{t\bar{t}}$ distribution at 8 TeV improves from $\chi^2/n_{\rm dat}$ =1.51 to 1.25, and the CMS $t\bar{t}b\bar{b}$ cross-section improves from 5.0 to 1.29. As expected in a global fit, given that the figure of merit being optimised is the total χ^2, eq. (4.4), for some datasets the overall fit quality is unchanged or slightly worsened as compared to the SM prediction.

From table 7, we notice that the only experiment for which the $\chi^2/n_{\rm dat}$ worsens significantly after the fit is the ATLAS $t\bar{t}Z$ cross-section measurement at 8 TeV, whose SM value of $\chi^2/n_{\rm dat} = 1.32$ increases to 5.29 after the fit. The origin of this poor χ^2 value can be traced back to some tension between the ATLAS and CMS measurements of the same observable. Indeed, as shown in figure 15, the ATLAS $t\bar{t}Z$ cross-section at 8 TeV lies somewhat below other measurements of the same quantity, in particular of the precise CMS measurement at 13 TeV. This exception aside, we find overall a good agreement between the theory calculations and the data used in the fit.

We now turn to present the comparisons between the results of the present SMEFT fit and the ATLAS and CMS input experimental data. We will also show comparisons for

JHEP04(2019)100

Dataset	$\chi^2/n_{\rm dat}$ (prior)	$\chi^2/n_{\rm dat}$ (fit)	$n_{\rm dat}$
ATLAS_tt_8TeV_ljets [$m_{t\bar{t}}$]	1.51	1.25	7
CMS_tt_8TeV_ljets [$y_{t\bar{t}}$]	1.17	1.17	10
CMS_tt2D_8TeV_dilep [$(m_{t\bar{t}}, y_t)$]	1.38	1.38	16
CMS_tt_13TeV_ljets2 [$m_{t\bar{t}}$]	1.09	1.28	8
CMS_tt_13TeV_dilep [$m_{t\bar{t}}$]	1.34	1.42	6
CMS_tt_13TeV_ljets_2016 [$m_{t\bar{t}}$]	1.87	1.87	10
ATLAS_WhelF_8TeV	1.98	0.27	3
CMS_WhelF_8TeV	0.31	1.18	3
CMS_ttbb_13TeV	5.00	1.29	1
CMS_tttt_13TeV	0.05	0.02	1
ATLAS_tth_13TeV	1.61	0.55	1
CMS_tth_13TeV	0.34	0.01	1
ATLAS_ttZ_8TeV	1.32	5.29	1
ATLAS_ttZ_13TeV	0.01	1.06	1
CMS_ttZ_8TeV	0.04	0.06	1
CMS_ttZ_13TeV	0.90	0.67	1
ATLAS_ttW_8TeV	1.34	0.27	1
ATLAS_ttW_13TeV	0.82	0.65	1
CMS_ttW_8TeV	1.54	0.54	1
CMS_ttW_13TeV	0.03	0.09	1
CMS_t_tch_8TeV_dif	0.11	0.32	6
ATLAS_t_tch_8TeV [y_t]	0.91	0.43	4
ATLAS_t_tch_8TeV [$y_{\bar{t}}$]	0.39	0.45	4
ATLAS_t_sch_8TeV	0.08	1.92	1
ATLAS_t_tch_13TeV	0.02	0.09	2
CMS_t_tch_13TeV_dif [y_t]	0.46	0.49	4
CMS_t_sch_8TeV	1.26	0.76	1
ATLAS_tW_inc_8TeV	0.02	0.06	1
CMS_tW_inc_8TeV	0.00	0.07	1
ATLAS_tW_inc_13TeV	0.52	0.82	1
CMS_tW_inc_13TeV	4.29	1.68	1
ATLAS_tZ_inc_13TeV	0.00	0.00	1
CMS_tZ_inc_13TeV	0.66	0.34	1
Total	**1.11**	**1.06**	**103**

Table 7. The values of the χ^2 per data point for each of the datasets included in the fit. In each case, we indicate the values of $\chi^2/n_{\rm dat}$ first when the theory calculations include only the SM contributions (second column) and then once they account for the SMEFT corrections, after the fit (third column). In the last column we indicate the number of data points $n_{\rm dat}$. Datasets are classified in three groups following the structure of tables 2–4: inclusive top quark pair production; $t\bar{t}$ production in association with heavy quarks, vector bosons, and Higgs bosons; and inclusive and associate production of single top quarks. In the case of datasets made of multiple differential distributions, we indicate the one that has been used in the analysis.

JHEP04(2019)100

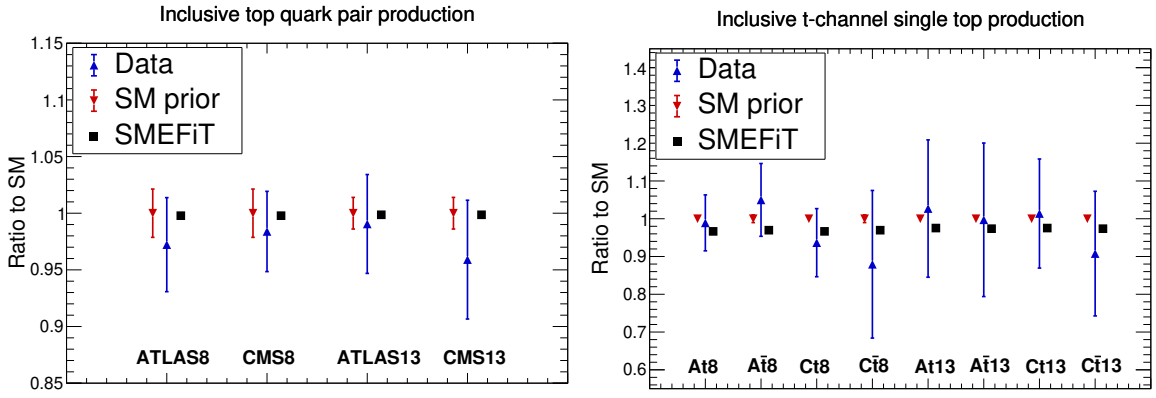

Figure 13. Comparison between ATLAS and CMS experimental data on the total inclusive $t\bar{t}$ (left) and single top t-channel (right) production cross-sections at 8 TeV and 13 TeV with the corresponding SM calculations and with the results of the SMEFT analysis. In the case of the SM calculations, we also shown the associated PDF uncertainties. Results are shown normalised to the central value of the SM prediction. Note that these inclusive cross-sections are not used as input to the fit (to avoid double counting with the corresponding differential distributions).

observables that are not included in the fit to avoid double counting, but which are anyway interesting to visualise in order to understand the main features of our results.

To begin with, in figure 13 we show a comparison between the ATLAS and CMS experimental data on the total inclusive $t\bar{t}$ and single top t-channel production cross-sections at 8 TeV and 13 TeV with the corresponding SM calculations and with the results of the SMEFT fit. For the single top case, we show separately the top and the anti-top cross-sections. In the case of the SM calculations, we also show the associated PDF uncertainties. Note that none of these total cross-sections (apart from the ATLAS 13 TeV single-top cross-section) are included in the SMEFT fit, since we already include the corresponding differential distributions. See section 3 for more details about measurements shown in this comparison. In figure 13, and in all subsequent comparisons, results are shown normalised to the central SM prediction.

From the comparisons in figure 13, we find good agreement between the data and the SM predictions. The SMEFT fit result typically moves towards the direction of the central experimental data point by an amount which corresponds to at most $|\delta_{\rm th}| \simeq 1\%$ and $\simeq 3\%$ of the SM prediction for inclusive $t\bar{t}$ and single-top production respectively, well below the experimental uncertainties. This SMEFT-induced shift in the theory predictions at the fit level is defined as

$$\delta_{\rm th} \equiv \frac{(\sigma_{\rm SMEFT} - \sigma_{\rm SM})}{\sigma_{\rm SM}} = \left(\sum_i^{N_{d6}} \kappa_i \frac{\langle c_i \rangle}{\Lambda^2} + \sum_{i,j}^{N_{d6}} \widetilde{\kappa}_{ij} \frac{\langle c_i c_j \rangle}{\Lambda^4} \right) \Big/ \sigma_{\rm SM} \,, \tag{5.1}$$

with $\langle c_i \rangle$ and $\langle c_i c_j \rangle$ represent the averages of the fitted SMEFT coefficients computed over the MC replica sample. While this shift is small for these precisely measured inclusive processes, this is not necessarily the case for differential distributions and for rarer top

JHEP04(2019)100

production processes, such as for single and top-pair production in association with vector bosons, as we will show below.

Next, in figure 14 we show a similar comparison as in figure 13 now for differential distributions in inclusive top quark pair and single top t-channel production. Specifically, we show the invariant mass distribution in $t\bar{t}$ production from ATLAS at 8 TeV and CMS at 13 TeV (2016 dataset), and the rapidity distributions in single top quark production in the t-channel from ATLAS at 8 TeV and from CMS at 13 TeV. In the latter case the top and anti-top quarks are combined into a single distribution.

From these comparisons, we find a similar level of agreement for the differential distributions as for the inclusive cross-sections. In the case of the $m_{t\bar{t}}$ distributions from ATLAS and CMS, the most marked effect comes from the rightmost bin of the distributions, where energy-growing effects are more important. We find that the SMEFT-induced shift is $\delta_{\rm th} = +13\%$ $(+40\%)$ at $m_{t\bar{t}} \simeq 1.4\,\text{TeV}$ (1.6 TeV) for the ATLAS 8 TeV (CMS 13 TeV) measurements. In section 5.4 we will show that results do not change if the $m_{t\bar{t}}$ distributions are replaced by the corresponding $y_{t\bar{t}}$ ones where the energy-growing effects are absent. In the case of the ATLAS $y_{t+\bar{t}}$ distribution in single top t-channel production, we observe how the data pulls the fit results. For this process, the SMEFT-induced shifts are around $\delta_{\rm th} \simeq -2.5\%$ for all the data bins for the rapidity distributions in t-channel single-top production, both at 8 TeV and at 13 TeV.

In figure 15 we show the corresponding comparison between experimental data and theory predictions for the ATLAS (labelled as 'A') and CMS (labelled as 'C') measurements of the cross-sections for single top production in the s-channel and in the tW and tZ associated production channels. We include in this comparison the results for the most updated measurements both at 8 TeV and at 13 TeV. In general, there is good agreement between the theory calculations and experimental data. The biggest SMEFT-induced shift is found for the s-channel cross-sections at 8 TeV, where $\delta_{\rm th} \simeq +35\%$. For single top production in association with a W boson, there is a negative shift of $\delta_{\rm th} \simeq -6\%$, similar for the two centre-of-mass energies. From the comparison of figure 15 we can also observe how in some cases the SMEFT fit interpolates between the ATLAS and CMS measurements, for instance for the $t + W$ cross-sections at 13 TeV and the s-channel cross-sections at 8 TeV.

Considering now the $t\bar{t}V$ processes, in figure 15 we show the corresponding plot for the measurements of the production cross-section of a top quark pair associated with a W or Z vector boson. We may observe here the origin of the poor agreement of the ATLAS $t\bar{t}Z$ measurement at 8 TeV with the theory prediction after the fit reported in table 7. Indeed, we find that for this process the ATLAS 8 TeV measurement (normalised to the SM prediction) barely agrees within uncertainties with the corresponding CMS 13 TeV cross-section, which exhibits the smallest uncertainties and thus dominates in the fit. For these $t\bar{t} + V$ processes, the SMEFT-induced shifts are $\delta_{\rm th} \simeq +23\%$ $(+11\%)$ for $t\bar{t} + W$ at 8 TeV (13 TeV) and $\delta_{\rm th} \simeq +26\%$ $(+31\%)$ for $t\bar{t} + Z$ at 8 TeV (13 TeV). These shifts are rather larger than for the corresponding inclusive cross-sections shown in figure 13, as allowed by the larger experimental uncertainties.

Finally, to complete this set of comparisons between the input experimental data and the corresponding theory calculations before and after the fit, we show in figure 16 the

JHEP04(2019)100

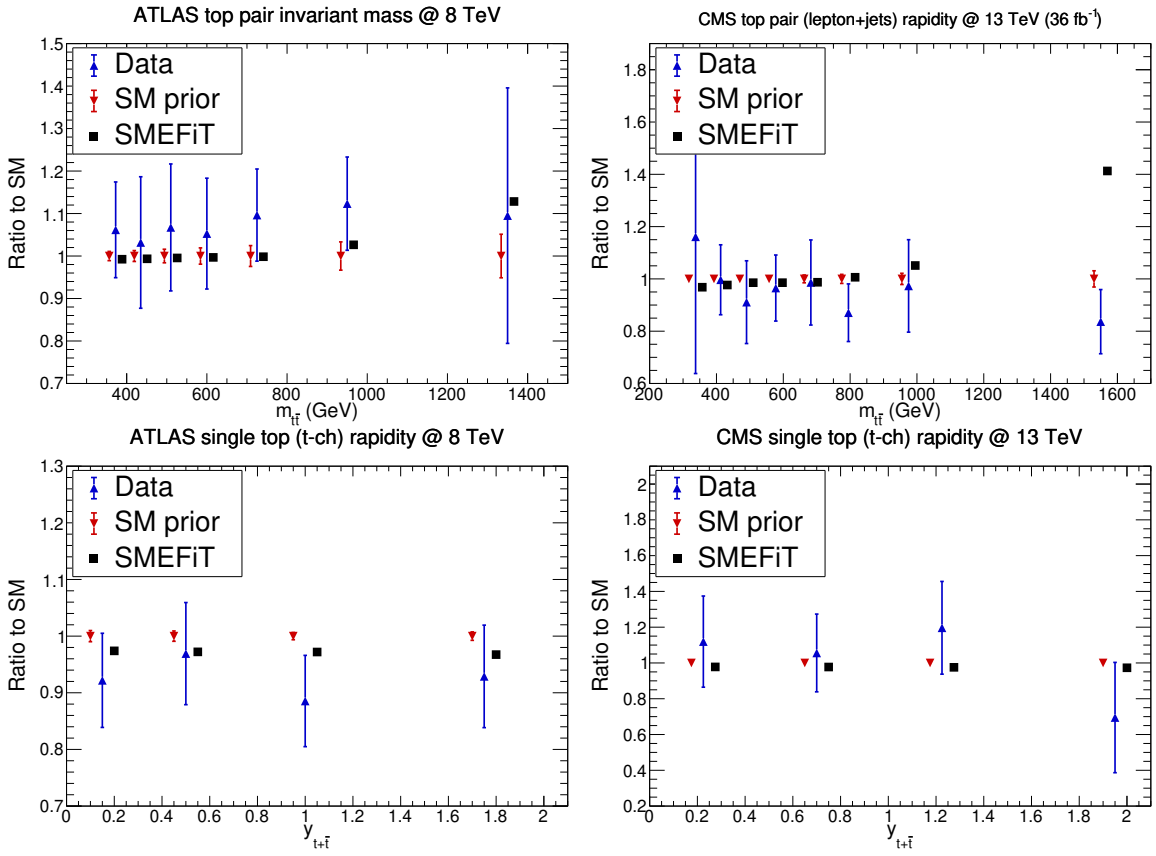

Figure 14. Same as figure 13, now for differential distributions in inclusive top quark pair and single top t-channel production. From left to right and from top to bottom we show the invariant mass distribution in $t\bar{t}$ production in ATLAS at 8 TeV, the rapidity distribution in $t\bar{t}$ production in CMS at 13 TeV (from the 2016 dataset based on $\mathcal{L} = 36\,\text{fb}^{-1}$), and the rapidity distributions for the $t + \bar{t}$ sum in single top quark production in the t-channel from ATLAS at 8 TeV and from CMS at 13 TeV.

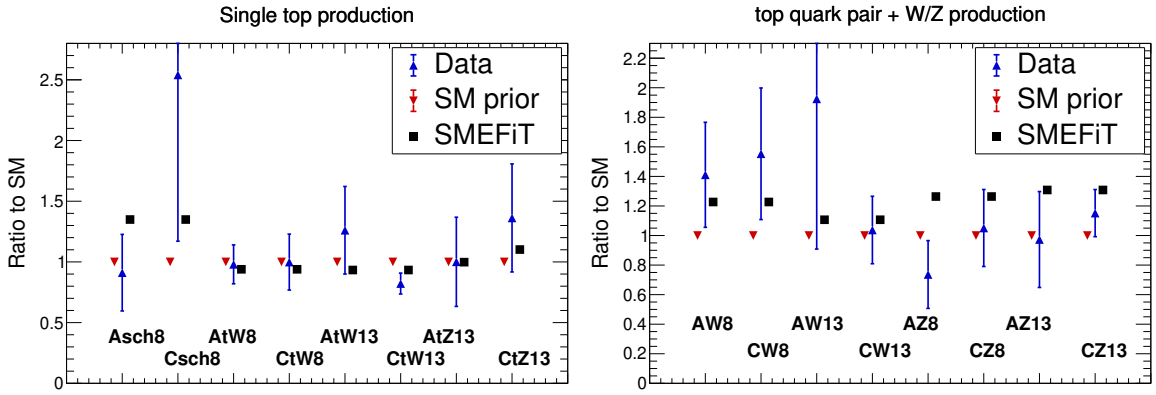

Figure 15. Left: the ATLAS (A) and CMS (C) measurements for single top production in the s-channel and in the tW and tZ associated production channel. Right: the corresponding comparison of the cross-section of $t\bar{t}$ associated production with W and Z vector bosons. We show the results for the most updated measurements both at 8 TeV and at 13 TeV.

JHEP04(2019)100

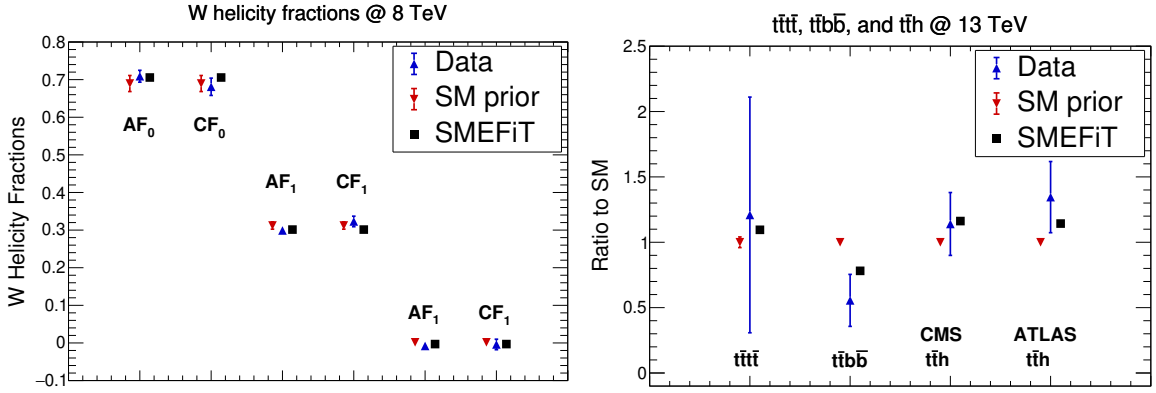

Figure 16. Left: same as figure 13 for the W helicity fractions F_0, F_1, and F_2 from ATLAS and CMS. Right: the corresponding comparison between data and theory predictions for the CMS measurements of $t\bar{t}b\bar{b}$ and $t\bar{t}t\bar{t}$ at 13 TeV, as well as for the $t\bar{t}h$ cross-section measurements from ATLAS and CMS at 13 TeV.

W helicity fractions F_0, F_1, and F_2 from ATLAS and CMS. There is good agreement between data and theory, and the $\delta_{\rm th}$ shifts are quite small. In the same figure, we also show the corresponding comparisons between data and theory predictions for the CMS measurements of $t\bar{t}b\bar{b}$ and $t\bar{t}t\bar{t}$ at 13 TeV, as well as for the $t\bar{t}H$ cross-section measurements from ATLAS and CMS at 13 TeV. Here the SMEFT-induced shifts are larger than for other processes, and we find $\delta_{\rm th} \simeq +10\%$ for $t\bar{t}t\bar{t}$ production, $\delta_{\rm th} \simeq -21\%$ for $t\bar{t}b\bar{b}$ production, and $\delta_{\rm th} \simeq +15\%$ for Higgs boson production in association with a $t\bar{t}$ pair.

As expected from the good agreement between the experimental data and the theory calculations already at the SM level reported in table 7, the overall pattern that is observed in these data/theory comparisons is that the SMEFT-induced shifts are (in relative terms) larger for observables with larger experimental uncertainties, and smaller for more precisely measured cross-sections such as in inclusive $t\bar{t}$ production. In all cases, these shifts $\delta_{\rm th}$ are smaller or at most comparable to the corresponding uncertainties of the experimental data.

5.2 The top quark degrees of freedom of the SMEFT

We now discuss the main results of this work. In the following, we present the fit results for the central values $\langle c_i \rangle$, eq. (4.5), and the corresponding 95% CL uncertainties, δc_i, for the $N_{\rm op} = 34$ dimension-6 SMEFT degrees of freedom relevant for the interpretation of top quark production measurements at the LHC. We also study the cross-correlations between these degrees of freedom. They provide an important piece of information since we know from the closure tests of section 4.4 that these correlations might be large because of flat directions in the parameter space.

In figure 17 we display the best-fit values of the $N_{\rm op} = 34$ degrees of freedom, $\langle c_i \rangle$, together with the corresponding 95% confidence levels δc_i. The dashed line indicates the SM prediction as reference. As with elsewhere in this work, we show the values of the degrees of freedom c_i/Λ^{-2} in units of TeV^{-2}, which coincide with c_i for $\Lambda = 1$ TeV. In the right panel of the same figure, we show the associated fit residuals r_i, eq. (4.10), which

JHEP04(2019)100

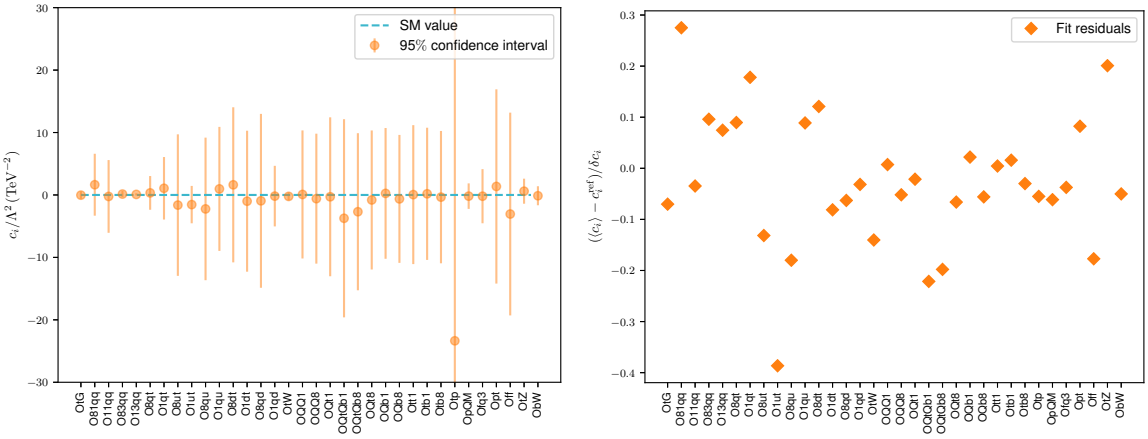

Figure 17. Left: the best-fit values, $\langle c_i \rangle$, with the corresponding 95% confidence intervals, δc_i, for the $N_{\rm op} = 34$ SMEFT degrees of freedom considered in this analysis. The dashed blue line indicates the SM, $c_i = 0$. Right plot: the associated fit residuals, which measure the deviation of the fit results with respect to the SM value $c_i^{(\rm ref)} = 0$ in units of the 95% CL uncertainties δc_i.

measure the deviation of the fit results with respect to the SM in units of the 95% CL δc_i uncertainties.

From this comparison, we find that the fit results are in good agreement with the SM within uncertainties, the fit residuals satisfying $|r_i| \leq 0.4$ for all operators. Note that the correlations between degrees of freedom imply that the fluctuations around the best-fit results are in general smaller as compared to the case in which all operators are completely independent.

From figure 17, we also observe that there is a rather wide range of values for the fit uncertainties δc_i obtained for the different degrees of freedom. For example, a very small uncertainty is found for the coefficients associated to `OtG` or `O83qq`, while much larger uncertainties are obtained for the fit coefficients associated to other degrees of freedom, including all the four-heavy-quark operators, such as `OQQ1`, and for `Otp`. In most cases, the origin of these differences in the size of the δc_i uncertainties can be traced back to table 6: different degrees of freedom are constrained by different processes, and in each case the available amount of experimental information varies widely. For instance, the four-heavy-quark operators are constrained by only two data points (the $b\bar{b}t\bar{t}$ and $t\bar{t}t\bar{t}$ cross-sections), hence the large uncertainties of the associated coefficients. Likewise, `Otp` is only constrained from the $t\bar{t}h$ cross-section measurements.

The interpretation of the 95% CL uncertainties shown in figure 17, requires some care. The reason is that the available data on top production at the LHC, summarised in tables 2–4, does not allow us to fully separate all possible independent directions in the SMEFT parameter space. As a consequence, as illustrated in section 4.4 at the closure test level, there will be in general large (anti-)correlations between the fit parameters, reflecting this degeneracy in the parameter space. As we will show now, in general more stringent bounds are obtained if each operator is fitted individually and the contributions of all other operators are set to zero, as compared to the bounds obtained in the global fit.

JHEP04(2019)100

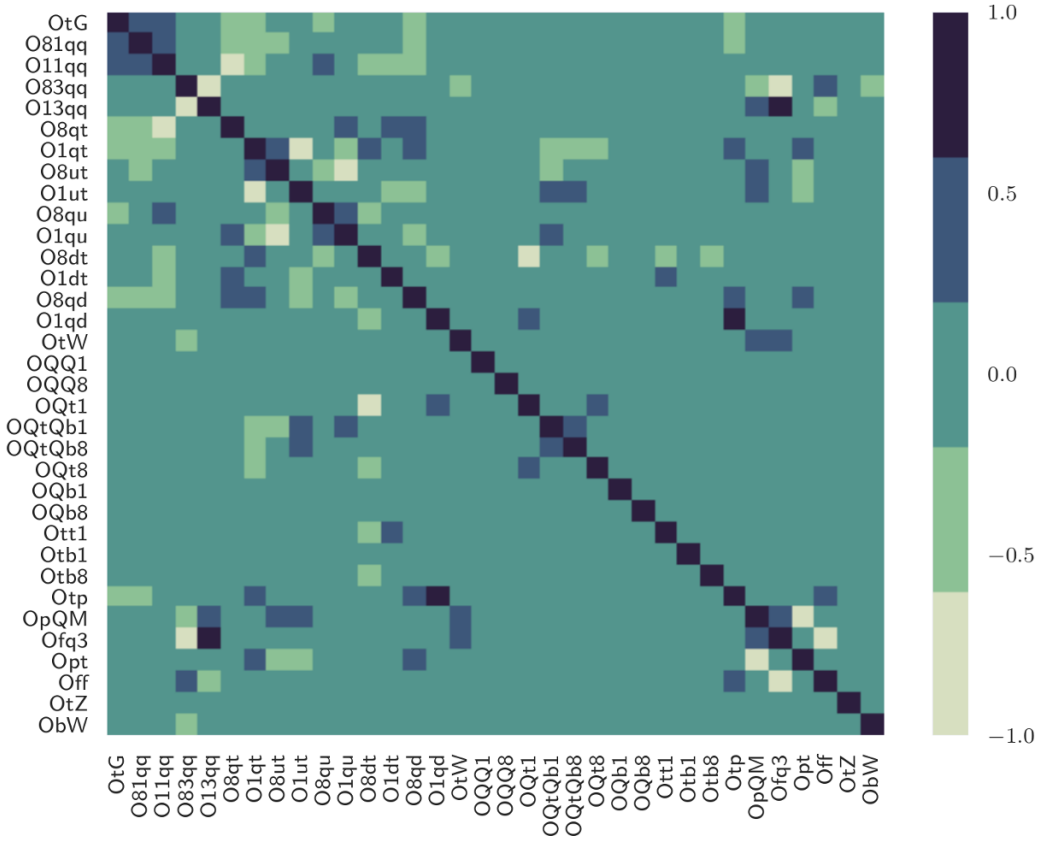

Figure 18. Heat map indicating the values of the correlation coefficient $\rho(c_i, c_j)$ between the $N_{\rm op} = 34$ fitted coefficients shown in figure 17, see text for more details.

To quantify this point, in figure 18 we show a heat map indicating the values of the correlation coefficient, eq. (4.11), between the 34 degrees of freedom constrained from the fit. In this heat map, dark blue regions correspond to degrees of freedom that are significantly correlated, while light green regions are instead degrees of freedom that are significantly anti-correlated. Indeed, we find that specific pairs of coefficients c_i exhibit a significant amount of (anti-)correlation, such as for instance `O1qd` and `Otp`. The effects of such correlations are ignored in fits where these degrees of freedom are constrained individually rather than marginalised from the global fit results, and lead in general to artificially tighter constraints.

Given the overall agreement between the fit results and the SM, it becomes possible to interpret the uncertainties δc_i as upper bounds on the parameter space of the SMEFT degrees of freedom. Such upper bounds provide important information for BSM model building, since they need to be satisfied for any UV-complete theory at high energies that has the SM as the low-energy effective theory. These bounds can also be compared with previous SMEFT studies of the top quark sector reported in the literature. While on the one hand our global SMEFT analysis is based on a wider LHC dataset than previous analysis of top quark production, on the other hand it explores a larger parameter space with reduced model assumptions. Therefore, a priori, one could either expect stronger (from the larger

JHEP04(2019)100

dataset) or weaker (from the reduction in model assumptions) bounds as compared to previous studies: only performing the actual fit itself can shed light on this question.

In order to compare with previous results, we will follow here the discussion in appendix A of the Top LHC Working Group EFT note [10], to where we direct the reader for further details. We note that the results quoted in [10] are in many cases restricted to fitting one operator at a time, or at most marginalising over a small subset of operators, and thus these limits might be too optimistic due to neglecting correlations with other directions in the SMEFT parameter space. We will quote here both the direct limits obtained from the top-quark measurements, and the indirect limits derived from non-top processes such as low-energy observables, the decays of B mesons, electroweak precision observables, and Drell-Yan production. See section 2.3 for a related discussion of the existing experimental constraints on SMEFT degrees of freedom that do not involve top quarks.

In table 8 we report the values of the 95% confidence level bounds (in units of TeV^{-2}, assuming $\Lambda = 1\,\text{TeV}$) for the coefficients of the 34 SMEFT degrees of freedom derived from the marginalisation of the results of the SMEFiT global analysis. We compare our results with those obtained elsewhere in the literature either from the direct analysis of top quark production ("direct") or from indirect bounds from other processes not involving top quarks ("indirect"). We note that for several degrees of freedom, such as for `Off` and `Otb1`, the bounds reported here have been obtained for the first time. In table 9 we additionally show the results for the differing theory settings used in the global fit; namely using only $\mathcal{O}(\Lambda^{-2})$ corrections and LO QCD in the SMEFT calculations.

As recommended in [10], it is important to also quote the bounds derived from fitting individual coefficients, one at a time, in order to compare them with the global fit results. The results from such single-operator fits are provided in table 10 using the same settings as in the baseline global fit (as well as by varying the theory settings, see the discussion in section 5.3). In the case of the individual fits of the operators that are very loosely constrained (in particular, for most of the four-heavy-quark degrees of freedom) we find that the SMEFiT approach is affected by numerical stability issues. Therefore, for such operators (identified in italics), it is more reliable to quote instead the 95% CL bounds obtained from the analytical minimisation of the χ^2, which for these cases has a relatively simple form.

By comparing the bounds obtained in the global and individual fits, tables 8, 9 and 10 respectively, one finds that for essentially all degrees of freedom the bounds obtained from the individual fits are either more stringent than or comparable to the marginalised results from the global fit. As discussed above, the reason for this can be traced back to the fact that within the single-operator fits one is neglecting cross-correlations between the different directions spanned by the fitted degrees of freedom. For instance, the 95% CL bound associated to `OtG` is $[-0.4, +0.4]$ in the global fit, while it is $[-0.08, +0.03]$ if the corresponding coefficient is fitted individually. Another example is `Otp`, whose bound is $[-60, +10]$ in the global fit, and $[-5.3, +1.6]$ in the individual fit, i.e. it is more stringent by about an order of magnitude.

Another important advantage of providing the results for the individual operators is that it allows us to better assess the impact that varying the theory settings has on the fit

JHEP04(2019)100

Notation	DoF	**SMEFiT** (marginalised)	Direct (other)	Indirect (other)
OQQ1	c_{QQ}^{1}	$[-9.4, 9.4]$		
OQQ8	c_{QQ}^{8}	$[-10, 9.4]$		
OQt1	c_{Qt}^{1}	$[-13, 12]$	$[-5.0, 4.9]$ [55]	
OQt8	c_{Qt}^{8}	$[-12, 10]$	$[-10.3, 9.3]$ [55]	
OQb1	c_{Qb}^{1}	$[-9.7, 9.7]$		
OQb8	c_{Qb}^{8}	$[-9.8, 9.3]$		
Ott1	c_{tt}^{1}	$[-11, 11]$	$[-2.9, 2.8]$ [55]	
Otb1	c_{tb}^{1}	$[-9.5, 9.9]$		
Otb8	c_{tb}^{8}	$[-10, 9.7]$		
OQtQb1	c_{QtQb}^{1}	$[-25, 8.9]$		
OQtQb8	c_{QtQb}^{8}	$[-18, 8.6]$		
O81qq	$c_{Qq}^{1,8}$	$[-4.7, 7.8]$	$[-6.9, 4.9]$ [55]	
O11qq	$c_{Qq}^{1,1}$	$[-6.8, 7.4]$	$[-3.1, 3.2]$ [55]	
O83qq	$c_{Qq}^{3,8}$	$[-1.3, 1.6]$	$[-6.1, 6.7]$ [55]	
O13qq	$c_{Qq}^{3,1}$	$[-1.1, 1.3]$	$[-0.7, 1.2]$ [48]	
O8qt	c_{tq}^{8}	$[-3.7, 4.1]$	$[-6.8, 3.5]$ [55]	
O1qt	c_{tq}^{1}	$[-5.3, 7.5]$	$[-2.8, 2.8]$ [55]	
O8ut	c_{tu}^{8}	$[-14, 10]$	$[-8.0, 4.8]$ [55]	
O1ut	c_{tu}^{1}	$[-5.8, 2.6]$	$[-3.6, 3.5]$ [55]	
O8qu	c_{Qu}^{8}	$[-15, 9.1]$	$[-8.1, 4.0]$ [55]	
O1qu	c_{Qu}^{1}	$[-9.8, 12]$	$[-3.3, 3.4]$ [55]	
O8dt	c_{td}^{8}	$[-9.5, 17]$	$[-12, 9.3]$ [55]	
O1dt	c_{td}^{1}	$[-13, 10]$	$[-4.9, 5.0]$ [55]	
O8qd	c_{Qd}^{8}	$[-14, 11]$	$[-11.8, 9.4]$ [55]	
O1qd	c_{Qd}^{1}	$[-3.5, 2.6]$	$[-5.0, 5.0]$ [55]	
OtG	c_{tG}	$[-0.4, 0.4]$	$[-1.3, 1.2]$ [48]	
OtW	c_{tW}	$[-1.8, 0.9]$	$[-4.0, 3.5]$ [48]	$[-2.8, 2.0]$ (EW)
ObW	c_{bW}	$[-2.6, 3.1]$		$[-15, 37]$ (EW)
OtZ	c_{tZ}	$[-2.1, 4.0]$	c_{tB}: $[-4.1, 4.1]$ [56]	c_{tB}: $[-5.8, 15.4]$ (EW)
Off	$c_{\varphi tb}$	$[-27, 8.7]$		
Ofq3	$c_{\varphi Q}^{3}$	$[-5.5, 5.8]$	$[-4.1, 2.0]$ [48]	
OpQM	$c_{\varphi Q}^{-}$	$[-3.5, 3]$	$c_{\varphi Q}^{1}$: $[-3.1, 3.1]$ [48]	$[-3.4, 7.4]$ (EW)
Opt	$c_{\varphi t}$	$[-13, 18]$	$[-9.7, 8, 3]$ [48]	$[-2.0, 5.6]$ (EW)
Otp	$c_{t\varphi}$	$[-60, 10]$		

Table 8. The 95% confidence level bounds (in units of TeV^{-2}, assuming $\Lambda = 1\,\text{TeV}$) for the coefficients of the 34 SMEFT degrees of freedom derived from the present analysis from the marginalisation of the global fit results. We also quote results obtained elsewhere from the direct analysis of top quark production and from indirect bounds from processes not involving top quarks.

JHEP04(2019)100

SMEFiT global analysis (marginalised)

Notation	DoF	Baseline	$\mathcal{O}(\Lambda^{-2})$ only	LO QCD
OQQ1	c_{QQ}^{1}	$[-9.4, 9.4]$	*	$[-9.6, 9.5]$
OQQ8	c_{QQ}^{8}	$[-10, 9.4]$	*	$[-9.9, 9.4]$
OQt1	c_{Qt}^{1}	$[-13, 12]$	*	$[-26, 12]$
OQt8	c_{Qt}^{8}	$[-12, 10]$	*	$[-20, 10]$
OQb1	c_{Qb}^{1}	$[-9.7, 9.7]$	*	$[-9.7, 9.8]$
OQb8	c_{Qb}^{8}	$[-9.8, 9.3]$	*	$[-9.7, 9.2]$
Ott1	c_{tt}^{1}	$[-11, 11]$	*	$[-17, 12]$
Otb1	c_{tb}^{1}	$[-9.5, 9.9]$	*	$[-9.6, 10]$
Otb8	c_{tb}^{8}	$[-10, 9.7]$	*	$[-11, 9.8]$
OQtQb1	c_{QtQb}^{1}	$[-25, 8.9]$	*	$[-19, 9.2]$
OQtQb8	c_{QtQb}^{8}	$[-18, 8.6]$	*	$[-15, 9.0]$
O81qq	$c_{Qq}^{1,8}$	$[-4.7, 7.8]$	$[-7.0, 12]$	$[-7.5, 7.6]$
O11qq	$c_{Qq}^{1,1}$	$[-6.8, 7.4]$	*	$[-17, 7.5]$
O83qq	$c_{Qq}^{3,8}$	$[-1.3, 1.6]$	$[-7.7, 9.1]$	$[-0.8, 1.3]$
O13qq	$c_{Qq}^{3,1}$	$[-1.1, 1.3]$	$[-0.5, 0.6]$	$[-1, 0.8]$
O8qt	c_{tq}^{8}	$[-3.7, 4.1]$	$[-10, 8.1]$	$[-3.2, 6.8]$
O1qt	c_{tq}^{1}	$[-5.3, 7.5]$	*	$[-6.1, 15]$
O8ut	c_{tu}^{8}	$[-14, 10]$	$[-13, 9.3]$	$[-22, 9.4]$
O1ut	c_{tu}^{1}	$[-5.8, 2.6]$	*	$[-8.8, 17]$
O8qu	c_{Qu}^{8}	$[-15, 9.1]$	$[-15, 8.5]$	$[-13, 16]$
O1qu	c_{Qu}^{1}	$[-9.8, 12]$	*	$[-14, 8.8]$
O8dt	c_{td}^{8}	$[-9.5, 17]$	$[-9.5, 13]$	$[-9.5, 45]$
O1dt	c_{td}^{1}	$[-13, 10]$	*	$[-18, 9.0]$
O8qd	c_{Qd}^{8}	$[-14, 11]$	$[-15, 11]$	$[-38, 16]$
O1qd	c_{Qd}^{1}	$[-3.5, 2.6]$	*	$[-1.9, 5.0]$
OtG	c_{tG}	$[-0.4, 0.4]$	$[-0.4, 0.4]$	$[-0.4, 0.4]$
OtW	c_{tW}	$[-1.8, 0.9]$	$[-0.8, 1.1]$	$[-2.0, 1.0]$
ObW	c_{bW}	$[-2.6, 3.1]$	*	$[-1.3, 4.8]$
OtZ	c_{tZ}	$[-2.1, 4.0]$	$[-14, 8.0]$	$[-4.6, 5.9]$
Off	$c_{\varphi tb}$	$[-27, 8.7]$	*	$[-8.9, 7.0]$
Ofq3	$c_{\varphi Q}^{3}$	$[-5.5, 5.8]$	$[-2.7, 2.2]$	$[-8.9, 7.0]$
OpQM	$c_{\varphi Q}^{-}$	$[-3.5, 3]$	$[-6.8, 11]$	$[-2.6, 3.3]$
Opt	$c_{\varphi t}$	$[-13, 18]$	$[-9.7, 20]$	$[-23, 7.3]$
Otp	$c_{t\varphi}$	$[-60, 10]$	$[-8.1, 5.0]$	$[-9.5, 11]$

Table 9. Same as table 8 for the 95% CL bounds obtained in the global fit, now comparing the results obtained using the baseline theory settings with those obtained when only the linear $\mathcal{O}(\Lambda^{-2})$ terms are included and when only LO QCD calculations are used for the SMEFT contribution.

JHEP04(2019)100

SMEFiT individual bounds (single-operator fits)

Notation	DoF	Baseline	$\mathcal{O}(\Lambda^{-2})$ only	LO QCD
OQQ1	c^1_{QQ}	$[-5.2, 4.9]$	$[-54, 83]$	$[-5.4, 5.2]$
OQQ8	c^8_{QQ}	$[-14, 12]$	$[-200, 18]$	$[-21, 16]$
OQt1	c^1_{Qt}	$[4.5, 4.5]$	$[-610, 210]$	$[-4.9, 4.9]$
OQt8	c^8_{Qt}	$[-10, 8.1]$	$[-69, 28]$	$[-11, 8.7]$
OQb1	c^1_{Qb}	$[6.9, 6.7]$	$[-1.9\ 10^3, -110]$	$[-6.1, 6.0]$
OQb8	c^8_{Qb}	$[-16, 12]$	$[-260, -14]$	$[-15, 11]$
Ott1	c^1_{tt}	$[-2.9, 2.7]$	$[-26, 41]$	$[-3.4, 3.2]$
Otb1	c^1_{tb}	$[-6.8, 6.8]$	$[-2.1\ 10^4, -1.4\ 10^3]$	$[-6.1, 6.1]$
Otb8	c^8_{tb}	$[-17, 12]$	$[-270, -15]$	$[-15, 11]$
OQtQb1	c^1_{QtQb}	$[-5.4, 5.5]$	$[160, 2.8\ 10^3]$	$[-4.8, 4.9]$
OQtQb8	c^8_{QtQb}	$[-14, 14]$	$[910, 1.6\ 10^4]$	$[-13, 13]$
O81qq	$c^{1,8}_{Qq}$	$[-0.6, 0.1]$	$[-1.2, 0.3]$	$[-0.6, 0.07]$
O11qq	$c^{1,1}_{Qq}$	$[-0.2, 0.02]$	$*$	$[-0.2, 0.03]$
O83qq	$c^{3,8}_{Qq}$	$[-0.5, 0.4]$	$[-3.3, -0.08]$	$[-0.7, 0.2]$
O13qq	$c^{3,1}_{Qq}$	$[-0.1, 0.09]$	$[-0.1, 0.2]$	$[-0.1, 0.09]$
O8qt	c^8_{tq}	$[-1.3, 0.4]$	$[-2.1, 1.5]$	$[-0.7, 0.09]$
O1qt	c^1_{tq}	$[-0.3, 0.02]$	$*$	$[-0.3, 0.03]$
O8ut	c^8_{tu}	$[-1.1, 0.04]$	$[-2.0, 0.09]$	$[-0.9, 0.03]$
O1ut	c^1_{tu}	$[-0.2, 0]$	$*$	$[-0.4, 0.03]$
O8qu	c^8_{Qu}	$[-2.6, 0.2]$	$[-4.4, 0.3]$	$[-2.6, 0.1]$
O1qu	c^1_{Qu}	$[-0.5, 0.02]$	$*$	$[-0.4, 0.03]$
O8dt	c^8_{td}	$[-2.5, -0.01]$	$[-4.6, -0.2]$	$[-1.6, 0.02]$
O1dt	c^1_{td}	$[-0.8, 0]$	$*$	$[-0.6, 0.03]$
O8qd	c^8_{Qd}	$[-2.7, 0.3]$	$[-3.7, 0.9]$	$[-1.9, 0.07]$
O1qd	c^1_{Qd}	$[-0.9, -0.01]$	$*$	$[-0.9, 0.05]$
OtG	c_{tG}	$[-0.08, 0.03]$	$[-0.08, 0.03]$	$[-0.1, 0.04]$
OtW	c_{tW}	$[-0.4, 0.2]$	$[-0.3, 0.1]$	$[-0.4, 0.2]$
ObW	c_{bW}	$[-0.6, 0.2]$	$*$	$[-0.7, 0.2]$
OtZ	c_{tZ}	$[-2.8, 4.5]$	$[-17, 4.6]$	$[-6.3, 7.4]$
Off	$c_{\varphi tb}$	$[-9.4, 9.5]$	$*$	$[-9.7, 9.8]$
Ofq3	$c^3_{\varphi Q}$	$[-0.9, 0.6]$	$[-1.0, 0.6]$	$[-1.0, 0.6]$
OpQM	$c^-_{\varphi Q}$	$[-4.2, 3.9]$	$[-4.2, 3.8]$	$[-5.1, 4.6]$
Opt	$c_{\varphi t}$	$[-6.4, 7.3]$	$[-6.9, 7.8]$	$[-7., 8.0]$
Otp	$c_{t\varphi}$	$[-5.3, 1.6]$	$[-5.1, 1.6]$	$[-5.4, 1.6]$

Table 10. Same as table 9, now for the results of individual fits when only one operator is constrained at a time. The bounds in italics have been obtained from the analytical minimisation of the χ^2 rather than using the SMEFiT numerical approach, see text for more details.

JHEP04(2019)100

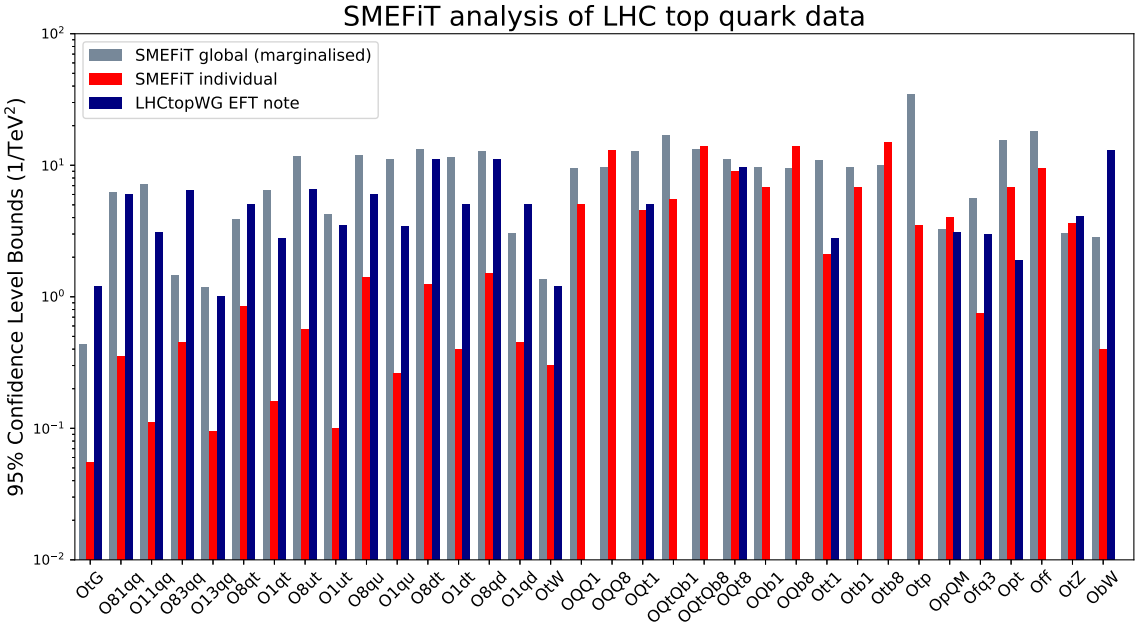

Figure 19. Graphical representation of the results of tables 8 and 10, where we compare the 95% CL bounds on the 34 degrees of freedom included the present analysis, both in the marginalised (global) and in the individual fit cases, with the bounds reported in the LHC Top WG EFT note [10].

results. For instance, as we will discuss in section 5.3, accounting for the quadratic $\mathcal{O}(\Lambda^{-4})$ terms leads to an improvement in the bounds of most operators, but assessing this effect is more transparently done in the case of the individual than in the global fits, where one has additional factors to take into account in the interpretation of the results.

The graphical representation of the comparison between the global fit results and the bounds reported in the LHC top WG EFT note (table 8), as well as with the individual fit results (table 10), is shown in figure 19. For the purposes of visualisation only, we have symmetrised the bounds reported there, that is, if a given operator has a 95% CL bound of $\left[\delta c_i^{(\min)}, \delta c_i^{(\max)}\right]$, then we show

$$\overline{\delta c_i} \equiv \left(\delta c_i^{(\max)} - \delta c_i^{(\min)}\right)/2\,. \tag{5.2}$$

We find that for some of the fitted degrees of freedom our bounds are stronger than those reported in previous studies, in some cases such as for `ObW` by nearly one order of magnitude. Another example is provided by the chromomagnetic operator `OtG`, for which the bound found in this work, $[-0.4, +0.4]$, is improved by a factor of three as compared to the bound quoted in the Top WG EFT note, $[-1.3, +1.2]$. From this comparison one can also appreciate how the individual bounds are in general rather tighter than the marginalised ones, except for some of the four-heavy-quark operators (and for `OtZ`) where they are instead comparable.

Another useful way to present our results is by representing the bounds on $\Lambda/\sqrt{|c_i|}$ that are derived from the fit. This is interesting because, assuming UV completions where

JHEP04(2019)100

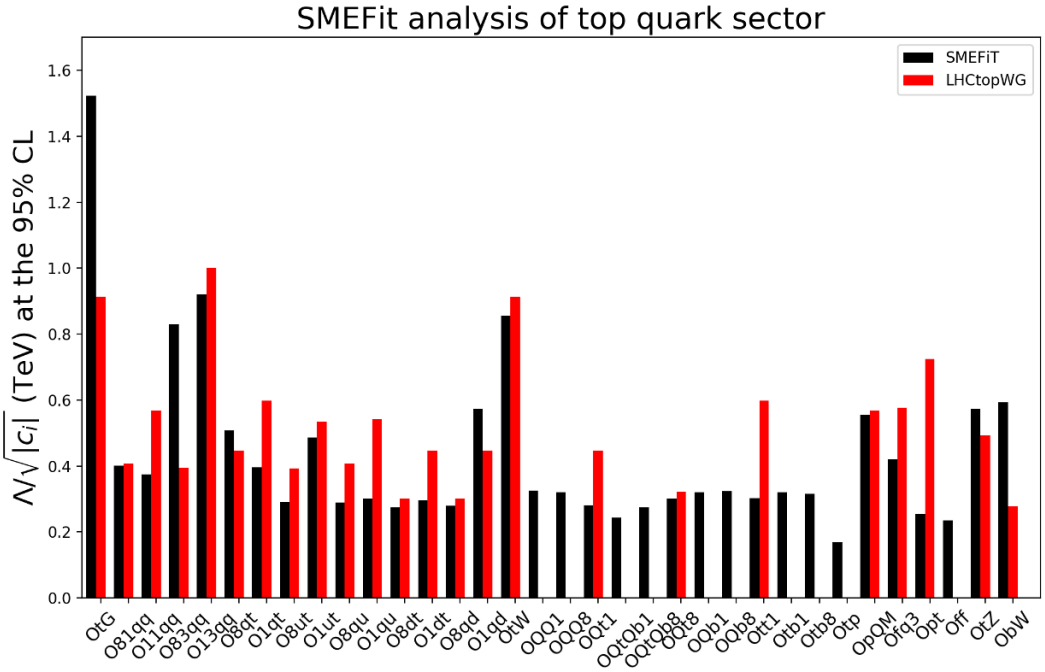

Figure 20. Same as figure 19, now representing the marginalised bounds obtained from the global fit as bounds on $\Lambda/\sqrt{|c_i|}$.

the values of the fitted degrees of freedom c_i are $\mathcal{O}(1)$, plotting the results this way indicates the approximate reach in energy that is being achieved by the SMEFT global analysis. This comparison is shown in figure 20, which is the analogous plot as figure 19 now representing the same bounds as bounds on the ratio $\Lambda/\sqrt{|c_i|}$ (now only for the marginalised bounds from the global fit). We find that for the degrees of freedom that are better constrained we achieve sensitivity up to scales as high as $\Lambda \simeq 1.5\,\text{TeV}$, in particular thanks to the chromomagnetic operator `OtG` which is well determined from the differential measurements of top quark pair production. Future measurements based on larger statistics should allow us to prove even higher scales, in particular by means of the high-luminosity LHC datasets.

5.3 The impact of the NLO QCD and $\mathcal{O}(\Lambda^{-4})$ corrections

The baseline fit results presented above are based on theory calculations that account both for the NLO QCD corrections to the SMEFT contributions and for the quadratic $\mathcal{O}(\Lambda^{-4})$ terms in eq. (2.2), see also the discussion in section 2. Here we aim to assess the robustness and stability of our results by comparing the baseline fit results with those of fits based on two alternative theory settings. Firstly we compare with a fit where only LO QCD effects are included for the SMEFT contributions, and then with a fit that includes only the linear $\mathcal{O}(\Lambda^{-2})$ terms in the effective theory expansion (but still based on NLO QCD for the SMEFT contributions).

These comparisons have been carried out in the case of both the marginalised results obtained from the global fit and of the fits to individual degrees of freedom. In table 9 we show the 95% CL bounds on the fitted degrees of freedom obtained in the global analysis,

JHEP04(2019)100

and compare the results obtained using the baseline theory settings with those obtained either when only the linear $\mathcal{O}(\Lambda^{-2})$ terms are included or when only LO QCD calculations are used for the SMEFT contribution. In table 10 we show the corresponding comparison in the case of individual fits. Recall that, as mentioned above, some of the individual bounds reported in table 10 have been evaluated from the analytical minimisation of the χ^2, which for those cases is more robust than the numerical minimisation.

As can be seen from table 10, the individual bounds that one obtains at $\mathcal{O}(\Lambda^{-2})$ are very loose for most of the four-heavy-quark operators. This indicates that, using only the linear SMEFT contribution, one has very limited sensitivity to these degrees of freedom. For this reason, we do not attempt to quote any bounds for the four-heavy-quark operators in the global fit based on $\mathcal{O}(\Lambda^{-2})$ theory in table 9: this small sensitivity might hinder the reliability of numerical approaches such as the ones we adopt here. This problem goes away once we include the $\mathcal{O}(\Lambda^{-4})$ contributions, due to the additional sensitivity provided by the quadratic terms. In this case, we can reliably quote 95% CL bounds for both global and individual fits.

In figure 21 we show the graphical representation of the bounds reported in table 9 for the global fit results with different theory settings. Note that, for the $\mathcal{O}(\Lambda^{-2})$ fit, several degrees of freedom are absent and we quote no bounds for the four-heavy-quark operators for the reasons mentioned above. For the case of the two global fits with theory variations, we do not show the comparison with the SM predictions or the fit residuals, as was done for the case of the fit with baseline theory settings in figure 17. The reason is that, in both cases, the agreement with the SM is as good as in the case of the global fit with baseline theory settings within uncertainties.

From these comparisons, one finds that the impact of the NLO QCD corrections varies depending on the specific operators considered. For the majority of operators, such as the two-light-two-heavy operators, the bounds derived from the data are either comparable or moderately improved once the NLO QCD corrections to the SMEFT contributions are accounted for. For example, the bound on the coefficient of the `O11qq` operator worsens from $[-6.8, +7.4]$, with baseline theory settings, to $[-17, +7.5]$, with LO QCD theory in the SMEFT contribution. Note, however, that this improvement does not necessarily mean that the NLO QCD corrections associated to the contributions of `O11qq` itself are important. Indeed, at the level of single-operator fits, the bounds at NLO and LO are quite similar, $[-0.2, 0.02]$ and $[-0.2, 0.03]$, respectively.

Similar considerations apply to those operators whose bounds in the global fit worsen when the NLO QCD corrections to the SMEFT contributions are missing. First, for these three operators, namely `Otp`, `Ofq3`, and `Off`, the bounds are relatively loose due to the limited fit sensitivity, so they are potentially affected by larger statistical fluctuations. Second, at the level of individual fits, one finds that including or not NLO QCD effects has essentially no impact on the resulting bounds. Therefore, the observed effect is most likely a consequence of the fact that adding NLO QCD corrections rearranges the weight of the different degrees of freedom in the global fit, leading to an overall modification of the bounds.

JHEP04(2019)100

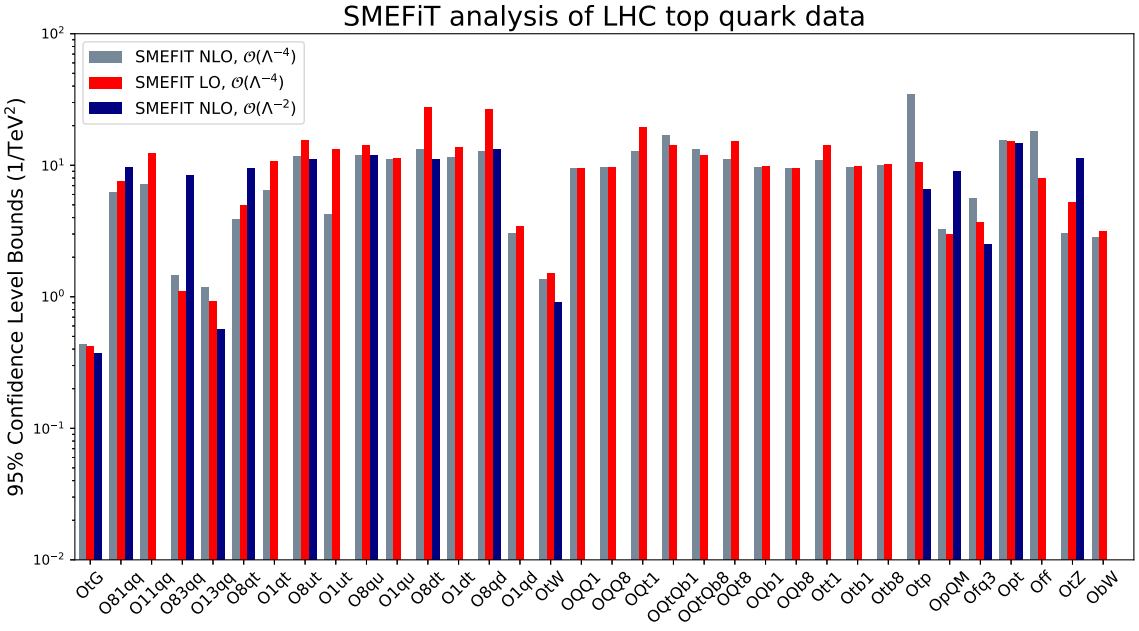

Figure 21. Same as figure 19, now comparing the baseline fit results reported in table 8 with those from table 9 based on the corresponding fits where (i) only LO QCD effects are included in the SMEFT contributions and (ii) only the linear $\mathcal{O}\left(\Lambda^{-2}\right)$ terms are accounted for (and thus neglecting the quadratic terms). Note that, at the level of the $\mathcal{O}\left(\Lambda^{-2}\right)$ fit, several degrees of freedom are absent and we quote no bounds for the four-heavy-quark operators.

Concerning the impact of the quadratic $\mathcal{O}\left(\Lambda^{-4}\right)$ terms, from the comparisons in figure 21, we find that for most degrees of freedom the bounds are similar regardless of whether or not these quadratic terms are included in the fit. This is the expected behaviour for those operators for which the dominant sensitivity in the fit arises already at $\mathcal{O}\left(\Lambda^{-2}\right)$, as indicated in table 6. For other operators, including $\mathcal{O}\left(\Lambda^{-4}\right)$ corrections leads to more stringent bounds. For instance, for the coefficient associated to the `O8qt` degree of freedom the baseline bounds of $[-3.7, +4.1]$ are degraded to $[-10, +8.1]$ if only the linear $\mathcal{O}\left(\Lambda^{-2}\right)$ corrections are taken into account. Another example is provided by `OtZ`, whose baseline bound of $[-2.1, +4.0]$ becomes rather looser in the linear approximation, $[-14, +8.0]$.

As already mentioned several times, within a global fit it is in general not possible to precisely pinpoint how a variation of the theory settings translates into a difference in the resulting constraints on the fitted degrees of freedom, with obvious exceptions such as for those operators whose contributions vanish at $\mathcal{O}\left(\Lambda^{-2}\right)$. For such assessment, the results of the single-operator fits reported in table 10 are more suitable. For example, from the results obtained in the single operator fits we can confirm that the improvement in the bounds obtained for the `O8qt` and `OtZ` degrees of freedom upon the inclusion of the quadratic $\mathcal{O}\left(\Lambda^{-4}\right]$ corrections is genuine, rather than an artefact of the global fit. The impact of including the quadratic terms is particularly manifest for the four-heavy-quark degrees of freedom, where one finds improvements of up to several orders of magnitude. For instance, while for `Otb1` the linear bounds are almost non-existent, $\left[-2 \cdot 10^4, -1.4 \cdot 10^3\right]$, they are improved down to $[-6.8, +6.8]$ once the $\mathcal{O}\left(\Lambda^{-4}\right)$ contributions are taken into account.

JHEP04(2019)100

Dataset	$\chi^2/n_{\rm dat}$			$n_{\rm dat}$
	NLO + $\mathcal{O}(\Lambda^{-4})$	NLO + $\mathcal{O}(\Lambda^{-2})$	LO + $\mathcal{O}(\Lambda^{-4})$	
ATLAS_tt_8TeV_ljets [$m_{t\bar{t}}$]	1.25	1.67	1.68	7
CMS_tt_8TeV_ljets [$y_{t\bar{t}}$]	1.17	1.04	0.82	10
CMS_tt2D_8TeV_dilep [$(m_{t\bar{t}}, y_t)$]	1.38	1.18	1.38	16
CMS_tt_13TeV_ljets2 [$m_{t\bar{t}}$]	1.28	0.67	0.67	8
CMS_tt_13TeV_dilep [$m_{t\bar{t}}$]	1.42	1.28	1.17	6
CMS_tt_13TeV_ljets_2016 [$m_{t\bar{t}}$]	1.87	1.17	0.57	10
ATLAS_WhelF_8TeV	0.27	0.77	1.97	3
CMS_WhelF_8TeV	1.18	0.25	1.43	3
CMS_ttbb_13TeV	1.29	3.02	1.14	1
CMS_tttt_13TeV	0.02	0.12	0.00	1
ATLAS_tth_13TeV	0.55	1.98	0.14	1
CMS_tth_13TeV	0.01	2.30	0.31	1
ATLAS_ttZ_8TeV	5.29	0.01	0.85	1
ATLAS_ttZ_13TeV	1.06	1.74	0.19	1
CMS_ttZ_8TeV	0.06	1.23	0.16	1
CMS_ttZ_13TeV	0.67	1.44	0.06	1
ATLAS_ttW_8TeV	0.27	2.08	0.12	1
ATLAS_ttW_13TeV	0.65	1.89	0.54	1
CMS_ttW_8TeV	0.54	1.96	0.35	1
CMS_ttW_13TeV	0.09	1.96	0.39	1
CMS_t_tch_8TeV_dif	0.32	0.46	0.24	6
ATLAS_t_tch_8TeV [y_t]	0.43	3.01	0.57	4
ATLAS_t_tch_8TeV [$y_{\bar{t}}$]	0.45	0.35	0.34	4
ATLAS_t_sch_8TeV	1.92	1.41	1.77	1
ATLAS_t_tch_13TeV	0.09	0.02	0.03	2
CMS_t_tch_13TeV_dif [y_t]	0.49	0.43	0.47	4
CMS_t_sch_8TeV	0.76	0.07	0.78	1
ATLAS_tW_inc_8TeV	0.06	0.00	0.14	1
CMS_tW_inc_8TeV	0.07	0.00	0.12	1
ATLAS_tW_inc_13TeV	0.82	0.57	0.91	1
CMS_tW_inc_13TeV	1.68	0.82	1.23	1
ATLAS_tZ_inc_13TeV	0.00	0.93	0.00	1
CMS_tZ_inc_13TeV	0.34	0.02	0.07	1
Total	**1.06**	**1.07**	**0.84**	**103**

Table 11. Same as table 7, now comparing the $\chi^2/n_{\rm dat}$ values of the fit obtained with the baseline theory settings (NLO QCD and $\mathcal{O}(\Lambda^{-4})$ corrections) with the corresponding values obtained in the fits where either only LO QCD effects or $\mathcal{O}(\Lambda^{-2})$ corrections are included.

To conclude this assessment of the impact of the NLO QCD and $\mathcal{O}(\Lambda^{-4})$ corrections on the fit results, in table 11 we show the comparison of the $\chi^2/n_{\rm dat}$ values of the fit obtained with the baseline theory settings (NLO QCD and $\mathcal{O}(\Lambda^{-4})$ corrections) with the corresponding values obtained in the fits where either only LO QCD effects or $\mathcal{O}(\Lambda^{-2})$ corrections are included in the SMEFT corrections. The corresponding comparison with the prior (SM) theory calculations was reported in table 7.

One finds that, for all three theory settings, the total $\chi^2/n_{\rm dat}$ is similar ($\simeq 1$). The lowest value is found when using LO QCD theory for the SMEFT corrections. In that case,

JHEP04(2019)100

we find $\chi^2/n_{\rm dat} = 0.84$, with the abundant data on $t\bar{t}$ differential distributions driving the improvement as compared to the baseline settings. We note that this effect is not statistically significant, so it could also be explained by a fluctuation. It will be interesting to revisit this comparison once more precise top production measurements become available, and assess whether or not there is evidence for the need of NLO QCD corrections to achieve the optimal description of the experimental data.

5.4 Dataset dependence and high-energy behaviour

Within the SMEFiT framework it is straightforward to repeat the analysis with arbitrary variations of the input dataset. To investigate the dependence of our results with respect to this choice of input dataset, in figure 22 we show a similar comparison as that of figure 19, now assessing how the baseline fit results vary if a different input dataset is used. In the first case, instead of the $m_{t\bar{t}}$ distributions indicated in table 7, we use the corresponding $y_{t\bar{t}}$ distributions for the inclusive $t\bar{t}$ production measurements. In the second case, the fit is performed only using inclusive $t\bar{t}$ production measurements as input, and excluding all other processes. Note that in the latter case the fit has sensitivity to only a subset of 15 degrees of freedom.

The rationale behind performing a fit replacing the $m_{t\bar{t}}$ distributions in inclusive top-quark pair production with the corresponding $y_{t\bar{t}}$ ones is to gauge the sensitivity of our results to the high-energy region, since the $m_{t\bar{t}}$ distribution is the one more directly sensitive to it. This was also illustrated by the large values of the SMEFT-induced shifts $\delta_{\rm th}$ found in the comparisons with experimental data at large $m_{t\bar{t}}$ in figure 14. Although high-energy measurements enhance the sensitivity to SMEFT effects, one should avoid being dominated by the highest energy bins since this could jeopardise the effective theory interpretation. Therefore, one would ideally like to see that the bounds do not become markedly worse once the $m_{t\bar{t}}$ distributions are replaced by the $y_{t\bar{t}}$ ones, since that would otherwise indicate that fit results are determined by high-energy events.

Concerning the fit based only on inclusive $t\bar{t}$ measurements, one would like to find that the bounds obtained from a SMEFT fit to a partial dataset are comparable to or looser than those from the baseline global dataset. Note that this is a non-trivial consistency check of the whole methodology; when additional experimental constraints are included in the analysis, then the bounds on the fitted coefficients must by necessity be either unchanged or smaller. If this were not the case, it would imply that fit results are driven not by the experimental data but by biased methodological choices.

From the comparison in figure 22 between the fits with either the baseline dataset or the $t\bar{t}$-only dataset, we find that the constraints on `OtG` are unchanged. This result is not unexpected, since it is well-known that the information on the chromomagnetic operator is dominated by inclusive $t\bar{t}$ production. We also observe that the bounds for some of the 2-light-2-heavy degrees of freedom such as `O83qq` and `O81qq` worsen, presumably as a consequence of the missing constraints provided by other processes, such as $t\bar{t}$ production in association with W or Z bosons. Indeed, for all the degrees of freedom directly constrained by the inclusive $t\bar{t}$ measurements, the bounds found in the global fit are comparable or superior to those obtained in the $t\bar{t}$-only fit.

JHEP04(2019)100

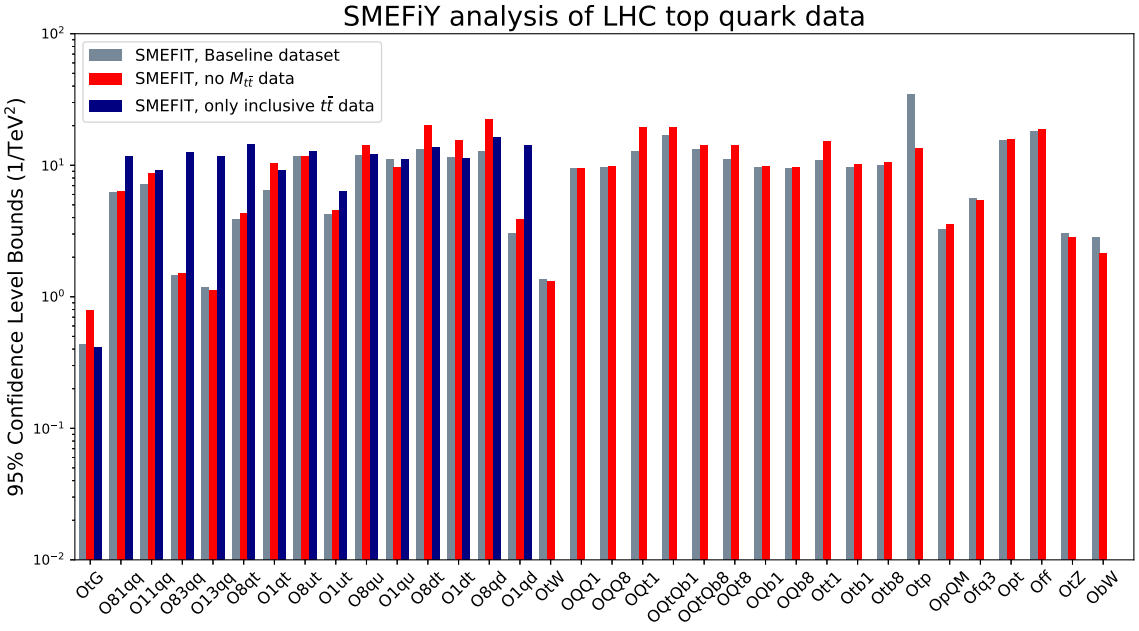

Figure 22. Same as figure 19, now comparing the fit results based on the baseline dataset with those obtained from two different datasets. In the first case, for the $t\bar{t}$ production measurements, instead of the $m_{t\bar{t}}$ distributions as indicated in table 7 we use the corresponding $y_{t\bar{t}}$ distributions. In the second case, the fit is performed only using inclusive $t\bar{t}$ production measurements as input.

The other comparison shown in figure 22 is that between the fit with the baseline dataset and with the same dataset where we have replaced the $m_{t\bar{t}}$ distributions with the corresponding $y_{t\bar{t}}$ ones. In this case, we find that the results are qualitatively stable, and do not display large differences. For a subset of the degrees of freedom, in particular those constrained by inclusive $t\bar{t}$ data, we find that somewhat more stringent bounds are obtained in the fits based on the $m_{t\bar{t}}$, rather than the $y_{t\bar{t}}$, distributions. For instance the bounds on the coefficient of `OtG` are found to be $[-0.4, +0.4]$ when fitting $m_{t\bar{t}}$ and $[-0.8, +0.8]$ when fitting instead $y_{t\bar{t}}$. These results suggest that indeed the fit benefits from the high-energy reach of the $m_{t\bar{t}}$ distributions, although only slightly.

Another way to study the impact that the SMEFT corrections have on the description of the experimental data at high energies is to focus on the constraints provided by the tail of the invariant mass distribution $m_{t\bar{t}}$ in top quark pair production, where energy-growing effects enhance the sensitivity to SMEFT corrections [165]. In order to highlight the impact that these energy-growing effects have on the description of the $m_{t\bar{t}}$ tails, it is useful to compute the shift induced by the SMEFT corrections to the SM calculation separated into the contributions from different degrees of freedom. For simplicity, in the following we restrict ourselves to the linear $\mathcal{O}\left(\Lambda^{-2}\right)$ corrections. In this case, following the notation of eq. (2.2), we want to compare the size of the individual corrections defined as

$$\Delta_i^{(\text{smeft})} \equiv \kappa_i \frac{\delta c_i}{\Lambda^2}\,, \qquad \widetilde{\Delta}_i^{(\text{smeft})} \equiv \widetilde{\kappa}_{ii} \frac{(\delta c_i)^2}{\Lambda^4}\,, \qquad i = 1, \ldots, N_{\text{op}}\,, \tag{5.3}$$

JHEP04(2019)100

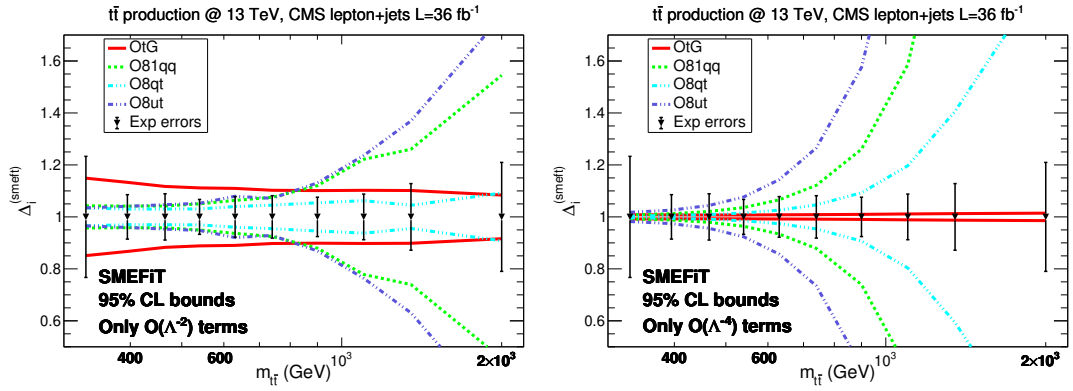

Figure 23. The shifts induced by representative SMEFT degrees of freedom to the SM cross-sections, eq. (5.3), for the $m_{t\bar{t}}$ distribution in the top quark pair production measurements at $\sqrt{s} = 13\,\text{TeV}$ from CMS, based on $\mathcal{L} = 36\,\text{fb}^{-1}$ and the lepton+jets channel [100]. We show the shifts arising from the linear (left) and from the purely quadratic (right) terms. The shifts $\Delta_i^{(\text{smeft})}$ have been computed assuming the 95% CL bounds δc_i of the baseline fit reported in table 8. For reference, we also indicate the size of the corresponding experimental uncertainties.

for the different bins of the $m_{t\bar{t}}$ distribution, and identify which degrees of freedom dominate at high energy. Note that, as discussed in section 2.5, in general there are several reasons why a given operator might or might not lead to energy-growing effects. In eq. (5.3), we will use as δc_i the 95% CL bounds for the baseline fit reported in table 8.

In figure 23 we show the values of the SMEFT-induced shifts, eq. (5.3), for the different bins of the $m_{t\bar{t}}$ distribution from the CMS measurement at 13 TeV in the lepton+jets final state, based on an integrated luminosity of $\mathcal{L} = 36\,\text{fb}^{-1}$ [100], which has the best coverage of the TeV region. To facilitate the visualisation, we restrict ourselves to the contributions associated to four representative degrees of freedom: `OtG`, `O81qq`, `O8qt`, and `O8ut`. For reference, we also show the corresponding total experimental uncertainty for each of the $m_{t\bar{t}}$ bins.

We observe that several operators lead to effects that grow with the energy. The steepest growth is found for the `O8ut` degree of freedom, but other operators that lead to energy-growing effects are `O81qq` and `O8qt`. Other operators are less sensitive to the high-energy region. This is illustrated by the case of `OtG`, whose sensitivity is concentrated in the $t\bar{t}$ threshold production region. It is therefore clear that pushing the reach of the experimental measurements deep into the TeV region will further increase the sensitivity to these energy-growing degrees of freedom. In this respect, a major concern will be to appropriately disentangle potential SMEFT signatures from the information used to constrain the proton structure in global fits, in particular the large-x gluon.

It should be emphasised that the individual shifts in figure 23 cannot be directly combined to construct the actual shift to the SM prediction in each cross-section bins, due to the replica-by-replica correlations between the various degrees of freedom. With this caveat, it is clear that the SMEFT-induced shifts could not be much larger than the bounds derived in this analysis without degrading the agreement between theory predictions and experimental data, a similar conclusion that what was derived from the comparisons with experimental data shown in figures 13–16.

JHEP04(2019)100

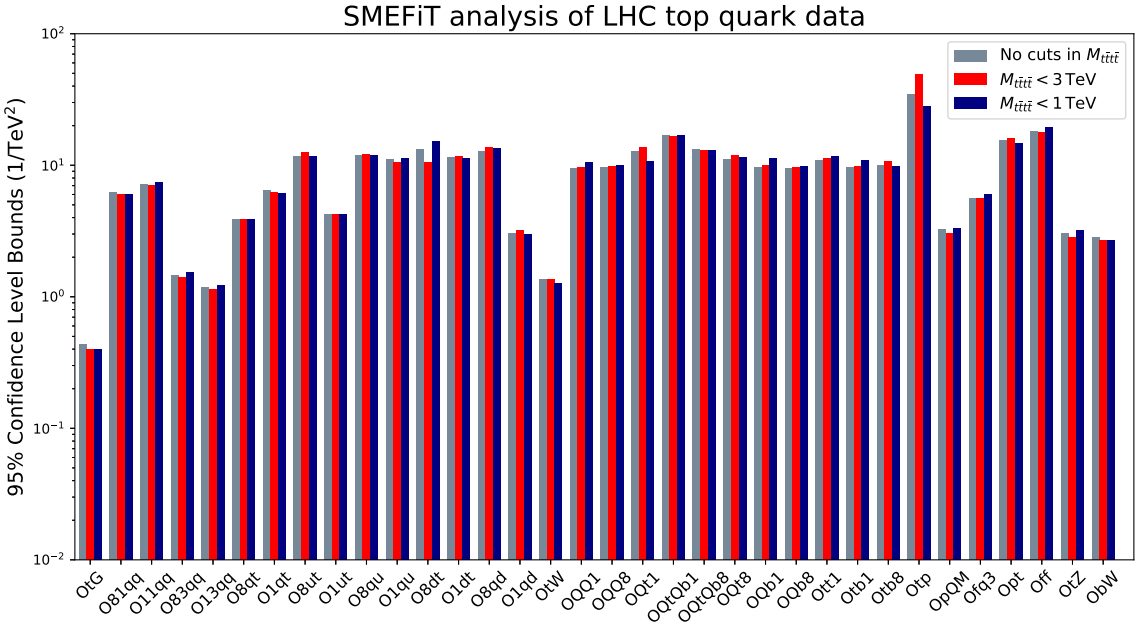

Figure 24. Same as figure 19, now comparing the baseline fit results with those obtained from two fits with differing cuts on $m_{t\bar{t}t\bar{t}}$ used in the theoretical calculation of the $t\bar{t}t\bar{t}$ production cross-section.

To conclude this discussion about high-energy effects, another of the input processes in the fit that in principle is sensitive to the high energy region is $t\bar{t}t\bar{t}$ production, where the invariant mass of the 4-top final state $m_{t\bar{t}t\bar{t}}$ can reach values of up to several TeV. In order to further assess the stability of our results with respect to the high-energy region, we have repeated the baseline fit imposing different cuts on the value of the 4-top invariant mass, from a loose cut requesting $m_{t\bar{t}t\bar{t}} \leq 3\,\text{TeV}$ to a more stringent cut with $m_{t\bar{t}t\bar{t}} \leq 1\,\text{TeV}$. The results of these fits are displayed in figure 24, and do not show any sensitivity to the value of $m_{t\bar{t}t\bar{t}}$ adopted in the theory calculation. We recall that in the current analysis a single $t\bar{t}t\bar{t}$ cross-section has been included; future measurements of this process, including possibly in differential form, could then become more sensitive to the high-energy region.

6 Summary and outlook

In this work we have presented a novel approach to carry out global analyses of the SMEFT. This new framework, which we have denoted by SMEFiT, is flexible, modular, robust upon enlarging the fitted parameter space, and resilient with respect to problems that arise frequently in SMEFT fits such as degeneracies and flat directions. Its main ingredients are the MC replica method to construct a representation of the probability distribution in the space of dimension-6 SMEFT degrees of freedom, and cross-validation, which prevents of over-fitting. Our results are provided as a sample of N_{rep} MC replicas, which can be used to derive predictions for related cross-sections and combined with other constraints on the SMEFT parameter space.

JHEP04(2019)100

As a proof-of-concept of the SMEFiT framework, we have presented a detailed analysis of top quark production measurements at the LHC 8 TeV and 13 TeV. We have included a wide range of top quark data, in terms of total rates and differential distributions. The theoretical SM and SMEFT cross-sections have been evaluated including NLO QCD corrections by default; in the SM case, we have also considered NNLO effects for the most accurately measured processes, namely differential distributions in $t\bar{t}$ and single top t-channel production. This combination of state-of-the art calculations with precision LHC measurements has allowed us to provide constraints on $N_{\rm op} = 34$ independent operators from the dimension-six Lagrangian in the Warsaw basis.

Our results are in good agreement with the SM expectations: we find that all the $N_{\rm op} = 34$ fitted SMEFT degrees of freedom are consistent with the SM result within uncertainties at the 95% CL. We have compared our results with existing bounds on the same operators presented in the literature, and have provided individual constraints on the operators in the SMEFiT framework. We have also studied the robustness of our results with respect to the inclusion of higher-order NLO QCD corrections, or $\mathcal{O}\left(\Lambda^{-4}\right)$ effects, and variations of the input dataset. We have found that including either NLO QCD corrections to the SMEFT contributions or the quadratic $\mathcal{O}\left(\Lambda^{-4}\right)$ terms leads to stronger bounds for most of the degrees of freedom in the fit.

The results of this analysis are available upon request as a sample of $N_{\rm rep} = 1000$ MC replicas representing the probability distribution in the space of Wilson coefficients for the $N_{\rm op} = 34$ SMEFT operators considered here. These replicas can be used to compute statistical properties of the distribution such as variances, correlations, and higher moments, and can be combined with other processes that provide complementary information on the SMEFT parameter space.

The study presented in this work is the first proof-of-principle application of the SMEFiT framework. Further studies and extensions can be envisioned. The next steps will be to consider a larger basis of fitted SMEFT operators by including other types of LHC processes beyond top quark production in the input dataset. These new measurements should include total rates and differential distributions in Higgs production, single and pair production of electroweak vector bosons, and also other processes directly sensitive to the TeV region, such as di-jet and multi-jet production. Eventually, one also might need to account for measurements from previous colliders such as LEP and from lower energy experiments. In this respect, our results pave the way towards a truly global fit of the SMEFT at dimension-six where direct constraints are simultaneously provided for the majority of the operators.

Acknowledgments

J.R. would like to thank Gerhard Raven and Wouter Verkerke for illuminating discussions. F.M. has received fundings from the European Union's Horizon 2020 research and innovation programme as part of the Marie Skłodowska-Curie Innovative Training Network MCnetITN3 (grant agreement no. 722104) and by the F.R.S.-FNRS under the 'Excellence of Science' EOS be.h project n. 30820817. Computational resources have been provided by

JHEP04(2019)100

the supercomputing facilities of the Université catholique de Louvain (CISM/UCL) and the Consortium des Équipements de Calcul Intensif en Fédération Wallonie Bruxelles (CÉCI). N.H., J.R., and E.S. are supported by the European Research Council Starting Grant "PDF4BSM". J.R. is also partially supported by the Netherlands Organization for Scientific Research (NWO). E.R.N. is supported by the European Commission through the Marie Skłodowska-Curie Action ParDHonS FFs.TMDs (grant number 752748), and was supported by the UK Science and Technology Facility Council through grant ST/P000630/1. E.V. is supported by a Marie Skłodowska-Curie Individual Fellowship of the European Commission's Horizon 2020 Programme under contract number 704187. C.Z. is supported by IHEP under Contract No. Y7515540U1.

JHEP04(2019)100

A Notation and conventions

In this appendix we summarise the notation and conventions that are adopted in this work concerning the relevant dimension-6 SMEFT operators. We adopt the notation of [3], where flavour indices are labelled by i, j, k and l; left-handed fermion SU(2) doublets are denoted by q, l; right-handed fermion singlets by u, d, e; the Higgs doublet by φ; the antisymmetric SU(2) tensor by $\varepsilon \equiv i\tau^2$; $\tilde{\varphi} = \varepsilon\varphi^*$; $(\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi) \equiv \varphi^\dagger(iD_\mu\varphi) - (iD_\mu\varphi^\dagger)\varphi$; $(\varphi^\dagger i\overleftrightarrow{D}^I_\mu \varphi) \equiv \varphi^\dagger\tau^I(iD_\mu\varphi) - (iD_\mu\varphi^\dagger)\tau^I\varphi$ where τ^I are the Pauli matrices; $T^A \equiv \lambda^A/2$ where λ^A are Gell-Mann matrices.

With these considerations, the dimension-6 SMEFT four-quark operators relevant for the interpretation of top quark measurements at the LHC are the following:

$$
\begin{aligned}
O_{qq}^{1(ijkl)} &= (\bar{q}_i\gamma^\mu q_j)(\bar{q}_k\gamma_\mu q_l),\\
O_{qq}^{3(ijkl)} &= (\bar{q}_i\gamma^\mu \tau^I q_j)(\bar{q}_k\gamma_\mu \tau^I q_l),\\
O_{qu}^{1(ijkl)} &= (\bar{q}_i\gamma^\mu q_j)(\bar{u}_k\gamma_\mu u_l),\\
O_{qu}^{8(ijkl)} &= (\bar{q}_i\gamma^\mu T^A q_j)(\bar{u}_k\gamma_\mu T^A u_l),\\
O_{qd}^{1(ijkl)} &= (\bar{q}_i\gamma^\mu q_j)(\bar{d}_k\gamma_\mu d_l),\\
O_{qd}^{8(ijkl)} &= (\bar{q}_i\gamma^\mu T^A q_j)(\bar{d}_k\gamma_\mu T^A d_l),\\
O_{uu}^{(ijkl)} &= (\bar{u}_i\gamma^\mu u_j)(\bar{u}_k\gamma_\mu u_l),\\
O_{ud}^{1(ijkl)} &= (\bar{u}_i\gamma^\mu u_j)(\bar{d}_k\gamma_\mu d_l),\\
O_{ud}^{8(ijkl)} &= (\bar{u}_i\gamma^\mu T^A u_j)(\bar{d}_k\gamma_\mu T^A d_l),\\
{}^\ddagger O_{quqd}^{1(ijkl)} &= (\bar{q}_i u_j)\,\varepsilon\,(\bar{q}_k d_l),\\
{}^\ddagger O_{quqd}^{8(ijkl)} &= (\bar{q}_i T^A u_j)\,\varepsilon\,(\bar{q}_k T^A d_l).
\end{aligned}
\tag{A.1}
$$

Recall that these operators satisfy all the symmetries of the SM, in particular gauge symmetry before electroweak symmetry breaking. Another class of relevant SMEFT operators are those that contain two quarks coupled to Higgs fields or gauge boson fields; the ones

relevant for top quark measurements are given by:

$$
\begin{aligned}
{}^{\ddagger}O_{u\varphi}^{(ij)} &= \bar{q}_i u_j \tilde{\varphi}\,(\varphi^\dagger\varphi), \\
O_{\varphi q}^{1(ij)} &= (\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi)(\bar{q}_i\gamma^\mu q_j), \\
O_{\varphi q}^{3(ij)} &= (\varphi^\dagger i\overleftrightarrow{D}^I_\mu \varphi)(\bar{q}_i\gamma^\mu \tau^I q_j), \\
O_{\varphi u}^{(ij)} &= (\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi)(\bar{u}_i\gamma^\mu u_j), \\
{}^{\ddagger}O_{\varphi ud}^{(ij)} &= (\tilde{\varphi}^\dagger i D_\mu \varphi)(\bar{u}_i\gamma^\mu d_j), \\
{}^{\ddagger}O_{uW}^{(ij)} &= (\bar{q}_i\sigma^{\mu\nu}\tau^I u_j)\,\tilde{\varphi}W^I_{\mu\nu}, \\
{}^{\ddagger}O_{dW}^{(ij)} &= (\bar{q}_i\sigma^{\mu\nu}\tau^I d_j)\,\varphi W^I_{\mu\nu}, \\
{}^{\ddagger}O_{uB}^{(ij)} &= (\bar{q}_i\sigma^{\mu\nu} u_j)\quad \tilde{\varphi}B_{\mu\nu}, \\
{}^{\ddagger}O_{uG}^{(ij)} &= (\bar{q}_i\sigma^{\mu\nu} T^A u_j)\,\tilde{\varphi}G^A_{\mu\nu}\,,
\end{aligned}
\tag{A.2}
$$

where $W^I_{\mu\nu}$ and $B_{\mu\nu}$ are the field-strength tensors of the electroweak interaction and $G^A_{\mu\nu}$ is the QCD one.

In eqs. (A.1) and (A.2), non-Hermitian operators are indicated with a double dagger symbol. In the case of Hermitian operators involving vector Lorentz bilinears, complex conjugation is the same as the transposition of generation indices: $O^{(ij)*} = O^{(ji)}$ and by extension, for four-fermion operators, $O^{(ijkl)*} = O^{(jilk)}$. In addition, it is understood in the notation above that the implicit sum over flavour indices only includes independent combinations.

References

[1] S. Weinberg, *Baryon and Lepton Nonconserving Processes*, *Phys. Rev. Lett.* **43** (1979) 1566 [INSPIRE].

[2] W. Buchmüller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys.* **B 268** (1986) 621 [INSPIRE].

[3] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085 [arXiv:1008.4884] [INSPIRE].

[4] S. Alioli, M. Farina, D. Pappadopulo and J.T. Ruderman, *Precision Probes of QCD at High Energies*, *JHEP* **07** (2017) 097 [arXiv:1706.03068] [INSPIRE].

[5] C. Englert, R. Kogler, H. Schulz and M. Spannowsky, *Higgs coupling measurements at the LHC*, *Eur. Phys. J.* **C 76** (2016) 393 [arXiv:1511.05170] [INSPIRE].

[6] J. Ellis, V. Sanz and T. You, *The Effective Standard Model after LHC Run I*, *JHEP* **03** (2015) 157 [arXiv:1410.7703] [INSPIRE].

[7] S. Alioli, M. Farina, D. Pappadopulo and J.T. Ruderman, *Catching a New Force by the Tail*, *Phys. Rev. Lett.* **120** (2018) 101801 [arXiv:1712.02347] [INSPIRE].

JHEP04(2019)100

[8] S. Alioli, W. Dekens, M. Girard and E. Mereghetti, *NLO QCD corrections to SM-EFT dilepton and electroweak Higgs boson production, matched to parton shower in POWHEG*, *JHEP* **08** (2018) 205 [arXiv:1804.07407] [INSPIRE].

[9] S. Alte, M. König and W. Shepherd, *Consistent Searches for SMEFT Effects in Non-Resonant Dijet Events*, *JHEP* **01** (2018) 094 [arXiv:1711.07484] [INSPIRE].

[10] D. Barducci et al., *Interpreting top-quark LHC measurements in the standard-model effective field theory*, arXiv:1802.07237 [INSPIRE].

[11] N. Castro, J. Erdmann, C. Grunwald, K. Kröninger and N.-A. Rosien, *EFTfitter — A tool for interpreting measurements in the context of effective field theories*, *Eur. Phys. J.* **C 76** (2016) 432 [arXiv:1605.05585] [INSPIRE].

[12] J. de Blas et al., *Electroweak precision observables and Higgs-boson signal strengths in the Standard Model and beyond: present and future*, *JHEP* **12** (2016) 135 [arXiv:1608.01509] [INSPIRE].

[13] J. Ellis, C.W. Murphy, V. Sanz and T. You, *Updated Global SMEFT Fit to Higgs, Diboson and Electroweak Data*, *JHEP* **06** (2018) 146 [arXiv:1803.03252] [INSPIRE].

[14] J. Ellis, V. Sanz and T. You, *Complete Higgs Sector Constraints on Dimension-6 Operators*, *JHEP* **07** (2014) 036 [arXiv:1404.3667] [INSPIRE].

[15] A. Buckley et al., *Results from TopFitter*, PoS(CKM2016)127 (2016) [arXiv:1612.02294] [INSPIRE].

[16] A. Butter, O.J.P. É boli, J. Gonzalez-Fraile, M.C. Gonzalez-Garcia, T. Plehn and M. Rauch, *The Gauge-Higgs Legacy of the LHC Run I*, *JHEP* **07** (2016) 152 [arXiv:1604.03105] [INSPIRE].

[17] A. Azatov, R. Contino, G. Panico and M. Son, *Effective field theory analysis of double Higgs boson production via gluon fusion*, *Phys. Rev.* **D 92** (2015) 035001 [arXiv:1502.00539] [INSPIRE].

[18] A. Biekötter, T. Corbett and T. Plehn, *The Gauge-Higgs Legacy of the LHC Run II*, arXiv:1812.07587 [INSPIRE].

[19] M. Schulze and Y. Soreq, *Pinning down electroweak dipole operators of the top quark*, *Eur. Phys. J.* **C 76** (2016) 466 [arXiv:1603.08911] [INSPIRE].

[20] V. Cirigliano, W. Dekens, J. de Vries and E. Mereghetti, *Constraining the top-Higgs sector of the Standard Model Effective Field Theory*, *Phys. Rev.* **D 94** (2016) 034031 [arXiv:1605.04311] [INSPIRE].

[21] S. Alioli, V. Cirigliano, W. Dekens, J. de Vries and E. Mereghetti, *Right-handed charged currents in the era of the Large Hadron Collider*, *JHEP* **05** (2017) 086 [arXiv:1703.04751] [INSPIRE].

[22] R. Franceschini, G. Panico, A. Pomarol, F. Riva and A. Wulzer, *Electroweak Precision Tests in High-Energy Diboson Processes*, *JHEP* **02** (2018) 111 [arXiv:1712.01310] [INSPIRE].

[23] J. Gao, L. Harland-Lang and J. Rojo, *The Structure of the Proton in the LHC Precision Era*, *Phys. Rept.* **742** (2018) 1 [arXiv:1709.04922] [INSPIRE].

[24] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, *J. Phys.* **G 43** (2016) 023001 [arXiv:1510.03865] [INSPIRE].

[25] J. Rojo et al., *The PDF4LHC report on PDFs and LHC data: Results from Run I and preparation for Run II*, *J. Phys.* **G 42** (2015) 103103 [arXiv:1507.00556] [INSPIRE].

JHEP04(2019)100

[26] NNPDF collaboration, *Unbiased determination of the proton structure function F_2^p with faithful uncertainty estimation*, *JHEP* **03** (2005) 080 [hep-ph/0501067] [INSPIRE].

[27] NNPDF collaboration, *Neural network determination of parton distributions: The Nonsinglet case*, *JHEP* **03** (2007) 039 [hep-ph/0701127] [INSPIRE].

[28] NNPDF collaboration, *A Determination of parton distributions with faithful uncertainty estimation*, *Nucl. Phys.* **B 809** (2009) 1 [*Erratum ibid.* **B 816** (2009) 293] [arXiv:0808.1231] [INSPIRE].

[29] NNPDF collaboration, *Precision determination of electroweak parameters and the strange content of the proton from neutrino deep-inelastic scattering*, *Nucl. Phys.* **B 823** (2009) 195 [arXiv:0906.1958] [INSPIRE].

[30] R.D. Ball et al., *A first unbiased global NLO determination of parton distributions and their uncertainties*, *Nucl. Phys.* **B 838** (2010) 136 [arXiv:1002.4407] [INSPIRE].

[31] R.D. Ball et al., *Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology*, *Nucl. Phys.* **B 849** (2011) 296 [arXiv:1101.1300] [INSPIRE].

[32] R.D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys.* **B 867** (2013) 244 [arXiv:1207.1303] [INSPIRE].

[33] NNPDF collaboration, *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040 [arXiv:1410.8849] [INSPIRE].

[34] NNPDF collaboration, *A Determination of the Charm Content of the Proton*, *Eur. Phys. J.* **C 76** (2016) 647 [arXiv:1605.06515] [INSPIRE].

[35] NNPDF collaboration, *A first unbiased global determination of polarized PDFs and their uncertainties*, *Nucl. Phys.* **B 887** (2014) 276 [arXiv:1406.5539] [INSPIRE].

[36] NNPDF collaboration, *Parton distributions from high-precision collider data*, *Eur. Phys. J.* **C 77** (2017) 663 [arXiv:1706.00428] [INSPIRE].

[37] NNPDF collaboration, *A determination of the fragmentation functions of pions, kaons and protons with faithful uncertainties*, *Eur. Phys. J.* **C 77** (2017) 516 [arXiv:1706.07049] [INSPIRE].

[38] NNPDF collaboration, *Charged hadron fragmentation functions from collider data*, *Eur. Phys. J.* **C 78** (2018) 651 [arXiv:1807.03310] [INSPIRE].

[39] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [arXiv:1405.0301] [INSPIRE].

[40] F. Maltoni, E. Vryonidou and C. Zhang, *Higgs production in association with a top-antitop pair in the Standard Model Effective Field Theory at NLO in QCD*, *JHEP* **10** (2016) 123 [arXiv:1607.05330] [INSPIRE].

[41] C. Degrande, F. Maltoni, K. Mimasu, E. Vryonidou and C. Zhang, *Single-top associated production with a Z or H boson at the LHC: the SMEFT interpretation*, *JHEP* **10** (2018) 005 [arXiv:1804.07773] [INSPIRE].

[42] M. Chala, J. Santiago and M. Spannowsky, *Constraining four-fermion operators using rare top decays*, arXiv:1809.09624 [INSPIRE].

[43] G. Durieux, F. Maltoni and C. Zhang, *Global approach to top-quark flavor-changing interactions*, *Phys. Rev.* **D 91** (2015) 074017 [arXiv:1412.7166] [INSPIRE].

JHEP04(2019)100

[44] J.A. Aguilar-Saavedra, *Effective four-fermion operators in top physics: A Roadmap*, *Nucl. Phys.* **B 843** (2011) 638 [*Erratum ibid.* **B 851** (2011) 443] [arXiv:1008.3562] [INSPIRE].

[45] J. D'Hondt, A. Mariotti, K. Mimasu, S. Moortgat and C. Zhang, *Learning to pinpoint effective operators at the LHC: a study of the* $t\bar{t}b\bar{b}$ *signature*, *JHEP* **11** (2018) 131 [arXiv:1807.02130] [INSPIRE].

[46] G. Durieux, J. Gu, E. Vryonidou and C. Zhang, *Probing top-quark couplings indirectly at Higgs factories*, *Chin. Phys.* **C 42** (2018) 123107 [arXiv:1809.03520] [INSPIRE].

[47] A. Buckley et al., *Global fit of top quark effective theory to data*, *Phys. Rev.* **D 92** (2015) 091501 [arXiv:1506.08845] [INSPIRE].

[48] A. Buckley et al., *Constraining top quark effective theory in the LHC Run II era*, *JHEP* **04** (2016) 015 [arXiv:1512.03360] [INSPIRE].

[49] S. Weinberg, *Phenomenological Lagrangians*, *Physica* **A 96** (1979) 327 [INSPIRE].

[50] C. Degrande et al., *Effective Field Theory: A Modern Approach to Anomalous Couplings*, *Annals Phys.* **335** (2013) 21 [arXiv:1205.4231] [INSPIRE].

[51] A. Kobach, *Baryon Number, Lepton Number and Operator Dimension in the Standard Model*, *Phys. Lett.* **B 758** (2016) 455 [arXiv:1604.05726] [INSPIRE].

[52] A. Falkowski, B. Fuks, K. Mawatari, K. Mimasu, F. Riva and V. Sanz, *Rosetta: an operator basis translator for Standard Model effective field theory*, *Eur. Phys. J.* **C 75** (2015) 583 [arXiv:1508.05895] [INSPIRE].

[53] R. Alonso, E.E. Jenkins, A.V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators III: Gauge Coupling Dependence and Phenomenology*, *JHEP* **04** (2014) 159 [arXiv:1312.2014] [INSPIRE].

[54] G. D'Ambrosio, G.F. Giudice, G. Isidori and A. Strumia, *Minimal flavor violation: An Effective field theory approach*, *Nucl. Phys.* **B 645** (2002) 155 [hep-ph/0207036] [INSPIRE].

[55] C. Zhang, *Constraining qqtt operators from four-top production: a case for enhanced EFT sensitivity*, *Chin. Phys.* **C 42** (2018) 023104 [arXiv:1708.05928] [INSPIRE].

[56] C. Englert and M. Russell, *Top quark electroweak couplings at future lepton colliders*, *Eur. Phys. J.* **C 77** (2017) 535 [arXiv:1704.01782] [INSPIRE].

[57] I. Brivio, Y. Jiang and M. Trott, *The SMEFTsim package, theory and tools*, *JHEP* **12** (2017) 070 [arXiv:1709.06492] [INSPIRE].

[58] P.L. Cho and E.H. Simmons, *Searching for G3 in* $t\bar{t}$ *production*, *Phys. Rev.* **D 51** (1995) 2360 [hep-ph/9408206] [INSPIRE].

[59] F. Krauss, S. Kuttimalai and T. Plehn, *LHC multijet events as a probe for anomalous dimension-six gluon interactions*, *Phys. Rev.* **D 95** (2017) 035024 [arXiv:1611.00767] [INSPIRE].

[60] V. Hirschi, F. Maltoni, I. Tsinikos and E. Vryonidou, *Constraining anomalous gluon self-interactions at the LHC: a reappraisal*, *JHEP* **07** (2018) 093 [arXiv:1806.04696] [INSPIRE].

[61] A. Falkowski and F. Riva, *Model-independent precision constraints on dimension-6 operators*, *JHEP* **02** (2015) 039 [arXiv:1411.0669] [INSPIRE].

[62] C. Grojean, W. Skiba and J. Terning, *Disguising the oblique parameters*, *Phys. Rev.* **D 73** (2006) 075008 [hep-ph/0602154] [INSPIRE].

JHEP04(2019)100

[63] I. Brivio and M. Trott, *Scheming in the SMEFT... and a reparameterization invariance!*, *JHEP* **07** (2017) 148 [*Addendum ibid.* **05** (2018) 136] [arXiv:1701.06424] [INSPIRE].

[64] K. Hagiwara, S. Ishihara, R. Szalapski and D. Zeppenfeld, *Low-energy effects of new interactions in the electroweak boson sector*, *Phys. Rev.* **D 48** (1993) 2182 [INSPIRE].

[65] J. Bramante, A. Delgado, L. Lehman and A. Martin, *Boosted Higgses from chromomagnetic b's: $b\bar{b}h$ at high luminosity*, *Phys. Rev.* **D 93** (2016) 053001 [arXiv:1410.3484] [INSPIRE].

[66] C. Zhang, *Single Top Production at Next-to-Leading Order in the Standard Model Effective Field Theory*, *Phys. Rev. Lett.* **116** (2016) 162002 [arXiv:1601.06163] [INSPIRE].

[67] C. Degrande, F. Maltoni, J. Wang and C. Zhang, *Automatic computations at next-to-leading order in QCD for top-quark flavor-changing neutral processes*, *Phys. Rev.* **D 91** (2015) 034024 [arXiv:1412.5594] [INSPIRE].

[68] D. Buarque Franzosi and C. Zhang, *Probing the top-quark chromomagnetic dipole moment at next-to-leading order in QCD*, *Phys. Rev.* **D 91** (2015) 114010 [arXiv:1503.08841] [INSPIRE].

[69] O. Bessidskaia Bylund, F. Maltoni, I. Tsinikos, E. Vryonidou and C. Zhang, *Probing top quark neutral couplings in the Standard Model Effective Field Theory at NLO in QCD*, *JHEP* **05** (2016) 052 [arXiv:1601.08193] [INSPIRE].

[70] G. Durieux, M. Perelló, M. Vos and C. Zhang, *Global and optimal probes for the top-quark effective field theory at future lepton colliders*, *JHEP* **10** (2018) 168 [arXiv:1807.02121] [INSPIRE].

[71] P. Artoisenet et al., *A framework for Higgs characterisation*, *JHEP* **11** (2013) 043 [arXiv:1306.6464] [INSPIRE].

[72] F. Maltoni, K. Mawatari and M. Zaro, *Higgs characterisation via vector-boson fusion and associated production: NLO and parton-shower effects*, *Eur. Phys. J.* **C 74** (2014) 2710 [arXiv:1311.1829] [INSPIRE].

[73] F. Demartin, F. Maltoni, K. Mawatari, B. Page and M. Zaro, *Higgs characterisation at NLO in QCD: CP properties of the top-quark Yukawa interaction*, *Eur. Phys. J.* **C 74** (2014) 3065 [arXiv:1407.5089] [INSPIRE].

[74] F. Demartin, F. Maltoni, K. Mawatari and M. Zaro, *Higgs production in association with a single top quark at the LHC*, *Eur. Phys. J.* **C 75** (2015) 267 [arXiv:1504.00611] [INSPIRE].

[75] K. Mimasu, V. Sanz and C. Williams, *Higher Order QCD predictions for Associated Higgs production with anomalous couplings to gauge bosons*, *JHEP* **08** (2016) 039 [arXiv:1512.02572] [INSPIRE].

[76] C. Degrande, B. Fuks, K. Mawatari, K. Mimasu and V. Sanz, *Electroweak Higgs boson production in the standard model effective field theory beyond leading order in QCD*, *Eur. Phys. J.* **C 77** (2017) 262 [arXiv:1609.04833] [INSPIRE].

[77] S. Fichet, A. Tonero and P. Rebello Teles, *Sharpening the shape analysis for higher-dimensional operator searches*, *Phys. Rev.* **D 96** (2017) 036003 [arXiv:1611.01165] [INSPIRE].

[78] C. Zhang, *Effective field theory approach to top-quark decay at next-to-leading order in QCD*, *Phys. Rev.* **D 90** (2014) 014008 [arXiv:1404.1264] [INSPIRE].

[79] E.E. Jenkins, A.V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators I: Formalism and lambda Dependence*, *JHEP* **10** (2013) 087 [arXiv:1308.2627] [INSPIRE].

JHEP04(2019)100

[80] E.E. Jenkins, A.V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators II: Yukawa Dependence*, *JHEP* **01** (2014) 035 [arXiv:1310.4838] [INSPIRE].

[81] N. Deutschmann, C. Duhr, F. Maltoni and E. Vryonidou, *Gluon-fusion Higgs production in the Standard Model Effective Field Theory*, *JHEP* **12** (2017) 063 [*Erratum ibid.* **02** (2018) 159] [arXiv:1708.00460] [INSPIRE].

[82] G. Panico, F. Riva and A. Wulzer, *Diboson Interference Resurrection*, *Phys. Lett.* **B 776** (2018) 473 [arXiv:1708.07823] [INSPIRE].

[83] A. Azatov, J. Elias-Miro, Y. Reyimuaji and E. Venturini, *Novel measurements of anomalous triple gauge couplings for the LHC*, *JHEP* **10** (2017) 027 [arXiv:1707.08060] [INSPIRE].

[84] A. Falkowski, M. Gonzalez-Alonso, A. Greljo, D. Marzocca and M. Son, *Anomalous Triple Gauge Couplings in the Effective Field Theory Approach at the LHC*, *JHEP* **02** (2017) 115 [arXiv:1609.06312] [INSPIRE].

[85] M. Farina, G. Panico, D. Pappadopulo, J.T. Ruderman, R. Torre and A. Wulzer, *Energy helps accuracy: electroweak precision tests at hadron colliders*, *Phys. Lett.* **B 772** (2017) 210 [arXiv:1609.08157] [INSPIRE].

[86] A. Greljo and D. Marzocca, *High-p_T dilepton tails and flavor physics*, *Eur. Phys. J.* **C 77** (2017) 548 [arXiv:1704.09015] [INSPIRE].

[87] A. Falkowski, M. González-Alonso and K. Mimouni, *Compilation of low-energy constraints on 4-fermion operators in the SMEFT*, *JHEP* **08** (2017) 123 [arXiv:1706.03783] [INSPIRE].

[88] R. Contino, A. Falkowski, F. Goertz, C. Grojean and F. Riva, *On the Validity of the Effective Field Theory Approach to SM Precision Tests*, *JHEP* **07** (2016) 144 [arXiv:1604.06444] [INSPIRE].

[89] O. Domenech, A. Pomarol and J. Serra, *Probing the SM with Dijets at the LHC*, *Phys. Rev.* **D 85** (2012) 074030 [arXiv:1201.6510] [INSPIRE].

[90] A. Biekötter, A. Knochel, M. Krämer, D. Liu and F. Riva, *Vices and virtues of Higgs effective field theories at large energy*, *Phys. Rev.* **D 91** (2015) 055029 [arXiv:1406.7320] [INSPIRE].

[91] A. Biekötter, J. Brehmer and T. Plehn, *Extending the limits of Higgs effective theory*, *Phys. Rev.* **D 94** (2016) 055032 [arXiv:1602.05202] [INSPIRE].

[92] A. Azatov, R. Contino, C.S. Machado and F. Riva, *Helicity selection rules and noninterference for BSM amplitudes*, *Phys. Rev.* **D 95** (2017) 065014 [arXiv:1607.05236] [INSPIRE].

[93] ATLAS collaboration, *Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *Eur. Phys. J.* **C 76** (2016) 538 [arXiv:1511.04716] [INSPIRE].

[94] CMS collaboration, *Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV*, *Eur. Phys. J.* **C 75** (2015) 542 [arXiv:1505.04480] [INSPIRE].

[95] CMS collaboration, *Measurement of double-differential cross sections for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV and impact on parton distribution functions*, *Eur. Phys. J.* **C 77** (2017) 459 [arXiv:1703.01630] [INSPIRE].

JHEP04(2019)100

[96] ATLAS collaboration, *Measurement of the W boson polarisation in $t\bar{t}$ events from pp collisions at $\sqrt{s} = 8$ TeV in the lepton + jets channel with ATLAS*, *Eur. Phys. J.* **C 77** (2017) 264 [*Erratum ibid.* **C 79** (2019) 19] [arXiv:1612.02577] [INSPIRE].

[97] CMS collaboration, *Measurement of the W boson helicity fractions in the decays of top quark pairs to lepton + jets final states produced in pp collisions at $\sqrt{s} = 8$ TeV*, *Phys. Lett.* **B 762** (2016) 512 [arXiv:1605.09047] [INSPIRE].

[98] CMS collaboration, *Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV*, *Phys. Rev.* **D 95** (2017) 092001 [arXiv:1610.04191] [INSPIRE].

[99] ATLAS collaboration, *Determination of the parton distribution functions of the proton from ATLAS measurements of differential W and Z/γ^* and $t\bar{t}$ cross sections*, ATL-PHYS-PUB-2018-017.

[100] CMS collaboration, *Measurement of differential cross sections for the production of top quark pairs and of additional jets in lepton+jets events from pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev.* **D 97** (2018) 112003 [arXiv:1803.08856] [INSPIRE].

[101] CMS collaboration, *Measurement of normalized differential $\mathrm{t\bar{t}}$ cross sections in the dilepton channel from pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **04** (2018) 060 [arXiv:1708.07638] [INSPIRE].

[102] CMS collaboration, *Measurements of $t\bar{t}$ cross sections in association with b jets and inclusive jets and their ratio using dilepton final states in pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett.* **B 776** (2018) 355 [arXiv:1705.10141] [INSPIRE].

[103] CMS collaboration, *Search for standard model production of four top quarks with same-sign and multilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J.* **C 78** (2018) 140 [arXiv:1710.10614] [INSPIRE].

[104] CMS collaboration, *Observation of top quark pairs produced in association with a vector boson in pp collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **01** (2016) 096 [arXiv:1510.01131] [INSPIRE].

[105] CMS collaboration, *Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **08** (2018) 011 [arXiv:1711.02547] [INSPIRE].

[106] ATLAS collaboration, *Measurement of the $t\bar{t}W$ and $t\bar{t}Z$ production cross sections in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *JHEP* **11** (2015) 172 [arXiv:1509.05276] [INSPIRE].

[107] ATLAS collaboration, *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ production cross sections in multilepton final states using 3.2 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J.* **C 77** (2017) 40 [arXiv:1609.01599] [INSPIRE].

[108] CMS collaboration, *Observation of $\mathrm{t\bar{t}}H$ production*, *Phys. Rev. Lett.* **120** (2018) 231801 [arXiv:1804.02610] [INSPIRE].

[109] ATLAS collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, *Phys. Lett.* **B 784** (2018) 173 [arXiv:1806.00425] [INSPIRE].

[110] CMS collaboration, *Measurement of the t-channel single-top-quark production cross section and of the $|V_{tb}|$ CKM matrix element in pp collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **06** (2014) 090 [arXiv:1403.7366] [INSPIRE].

JHEP04(2019)100

[111] CMS collaboration, *Search for s channel single top quark production in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, *JHEP* **09** (2016) 027 [arXiv:1603.02555] [INSPIRE].

[112] ATLAS collaboration, *Evidence for single top-quark production in the s-channel in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector using the Matrix Element Method*, *Phys. Lett.* **B 756** (2016) 228 [arXiv:1511.05980] [INSPIRE].

[113] ATLAS collaboration, *Fiducial, total and differential cross-section measurements of t-channel single top-quark production in pp collisions at 8 TeV using data collected by the ATLAS detector*, *Eur. Phys. J.* **C 77** (2017) 531 [arXiv:1702.02859] [INSPIRE].

[114] ATLAS collaboration, *Measurement of the inclusive cross-sections of single top-quark and top-antiquark t-channel production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **04** (2017) 086 [arXiv:1609.03920] [INSPIRE].

[115] CMS collaboration, *Cross section measurement of t-channel single top quark production in pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett.* **B 772** (2017) 752 [arXiv:1610.00678] [INSPIRE].

[116] CMS collaboration, *Single top t-channel differential cross section at 8 TeV*, CMS-PAS-TOP-14-004.

[117] CMS collaboration, *Measurement of the differential cross section for t-channel single-top-quark production at $\sqrt{s} = 13$ TeV*, CMS-PAS-TOP-16-004.

[118] ATLAS collaboration, *Measurement of the production cross-section of a single top quark in association with a W boson at 8 TeV with the ATLAS experiment*, *JHEP* **01** (2016) 064 [arXiv:1510.03752] [INSPIRE].

[119] CMS collaboration, *Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s} = 8$ TeV*, *Phys. Rev. Lett.* **112** (2014) 231802 [arXiv:1401.2942] [INSPIRE].

[120] ATLAS collaboration, *Measurement of the cross-section for producing a W boson in association with a single top quark in pp collisions at $\sqrt{s} = 13$ TeV with ATLAS*, *JHEP* **01** (2018) 063 [arXiv:1612.07231] [INSPIRE].

[121] CMS collaboration, *Measurement of the production cross section for single top quarks in association with W bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **10** (2018) 117 [arXiv:1805.07399] [INSPIRE].

[122] CMS collaboration, *Measurement of the associated production of a single top quark and a Z boson in pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett.* **B 779** (2018) 358 [arXiv:1712.02825] [INSPIRE].

[123] ATLAS collaboration, *Measurement of the production cross-section of a single top quark in association with a Z boson in proton-proton collisions at 13 TeV with the ATLAS detector*, *Phys. Lett.* **B 780** (2018) 557 [arXiv:1710.03659] [INSPIRE].

[124] M. Czakon, M.L. Mangano, A. Mitov and J. Rojo, *Constraints on the gluon PDF from top quark pair production at hadron colliders*, *JHEP* **07** (2013) 167 [arXiv:1303.7215] [INSPIRE].

[125] M. Czakon, D. Heymes and A. Mitov, *Dynamical scales for multi-TeV top-pair production at the LHC*, *JHEP* **04** (2017) 071 [arXiv:1606.03350] [INSPIRE].

[126] J. Gao and A.S. Papanastasiou, *Top-quark pair-production and decay at high precision*, *Phys. Rev.* **D 96** (2017) 051501 [arXiv:1705.08903] [INSPIRE].

[127] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007 [arXiv:0811.4622] [INSPIRE].

JHEP04(2019)100

[128] R. Boughezal et al., *Color singlet production at NNLO in MCFM*, *Eur. Phys. J.* **C 77** (2017) 7 [arXiv:1605.08011] [INSPIRE].

[129] M. Czakon, D. Heymes and A. Mitov, *fastNLO tables for NNLO top-quark pair differential distributions*, arXiv:1704.08551 [INSPIRE].

[130] M. Czakon, N.P. Hartland, A. Mitov, E.R. Nocera and J. Rojo, *Pinning down the large-x gluon with NNLO top-quark pair differential distributions*, *JHEP* **04** (2017) 044 [arXiv:1611.08609] [INSPIRE].

[131] CMS collaboration, *Measurements of* $\mathrm{t\bar{t}}$ *differential cross sections in proton-proton collisions at* $\sqrt{s} = 13$ *TeV using events containing two leptons*, *JHEP* **02** (2019) 149 [arXiv:1811.06625] [INSPIRE].

[132] ATLAS collaboration, *Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at* $\sqrt{s} = 13$ *TeV using the ATLAS detector*, *JHEP* **11** (2017) 191 [arXiv:1708.00727] [INSPIRE].

[133] ATLAS collaboration, *Measurements of top-quark pair differential cross-sections in the* $e\mu$ *channel in pp collisions at* $\sqrt{s} = 13$ *TeV using the ATLAS detector*, *Eur. Phys. J.* **C 77** (2017) 292 [arXiv:1612.05220] [INSPIRE].

[134] ATLAS collaboration, *Measurements of* $t\bar{t}$ *differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in pp collisions at* $\sqrt{s} = 13$ *TeV using the ATLAS detector*, *Phys. Rev.* **D 98** (2018) 012003 [arXiv:1801.02052] [INSPIRE].

[135] M. Czakon, D. Heymes, A. Mitov, D. Pagani, I. Tsinikos and M. Zaro, *Top-pair production at the LHC through NNLO QCD and NLO EW*, *JHEP* **10** (2017) 186 [arXiv:1705.04105] [INSPIRE].

[136] M. Czakon et al., *Resummation for (boosted) top-quark pair production at NNLO+NNLL' in QCD*, *JHEP* **05** (2018) 149 [arXiv:1803.07623] [INSPIRE].

[137] ATLAS collaboration, *Combination of the ATLAS and CMS measurements of the W-boson polarization in top-quark decays*, ATLAS-CONF-2013-033.

[138] CMS collaboration, *Measurements of* $t\bar{t}$ *spin correlations and top quark polarization using dilepton final states in pp collisions at* $\sqrt{s} = 8$ *TeV*, *Phys. Rev.* **D 93** (2016) 052007 [arXiv:1601.01107] [INSPIRE].

[139] ATLAS collaboration, *Measurements of top-quark pair spin correlations in the* $e\mu$ *channel at* $\sqrt{s} = 13$ *TeV using pp collisions in the ATLAS detector*, ATLAS-CONF-2018-027.

[140] ATLAS collaboration, *Measurement of the* $t\bar{t}\gamma$ *production cross section in proton-proton collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, *JHEP* **11** (2017) 086 [arXiv:1706.03046] [INSPIRE].

[141] ATLAS collaboration, *Observation of top-quark pair production in association with a photon and measurement of the* $t\bar{t}\gamma$ *production cross section in pp collisions at* $\sqrt{s} = 7$ *TeV using the ATLAS detector*, *Phys. Rev.* **D 91** (2015) 072007 [arXiv:1502.00586] [INSPIRE].

[142] CMS collaboration, *Measurement of the semileptonic* $\mathrm{t\bar{t}} + \gamma$ *production cross section in pp collisions at* $\sqrt{s} = 8$ *TeV*, *JHEP* **10** (2017) 006 [arXiv:1706.08128] [INSPIRE].

[143] CMS collaboration, *Measurement of* $\mathrm{t\bar{t}}$ *production with additional jet activity, including* b *quark jets, in the dilepton decay channel using pp collisions at* $\sqrt{s} = 8$ *TeV*, *Eur. Phys. J.* **C 76** (2016) 379 [arXiv:1510.03072] [INSPIRE].

[144] C. Englert, M. Russell and C.D. White, *Effective Field Theory in the top sector: do multijets help?*, *Phys. Rev.* **D 99** (2019) 035019 [arXiv:1809.09744] [INSPIRE].

JHEP04(2019)100

[145] ATLAS collaboration, *Measurements of fiducial and differential cross-sections of $t\bar{t}$ production with additional heavy-flavour jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, ATLAS-CONF-2018-029.

[146] ATLAS collaboration, *Measurements of fiducial cross-sections for $t\bar{t}$ production with one or two additional b-jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *Eur. Phys. J.* **C 76** (2016) 11 [arXiv:1508.06868] [INSPIRE].

[147] CMS collaboration, *Search for Standard Model Production of Four Top Quarks in the Lepton + Jets Channel in pp Collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **11** (2014) 154 [arXiv:1409.7339] [INSPIRE].

[148] CMS collaboration, *Search for standard model production of four top quarks in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett.* **B 772** (2017) 336 [arXiv:1702.06164] [INSPIRE].

[149] ATLAS collaboration, *Search for four-top-quark production in final states with one charged lepton and multiple jets using 3.2 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC*, ATLAS-CONF-2016-020.

[150] ATLAS collaboration, *Search for four-top-quark production in the single-lepton and opposite-sign dilepton final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev.* **D 99** (2019) 052009 [arXiv:1811.02305] [INSPIRE].

[151] A. Giammanco, *Single top quark production at the LHC*, *Rev. Phys.* **1** (2016) 1 [arXiv:1511.06748] [INSPIRE].

[152] E.L. Berger, J. Gao, C.P. Yuan and H.X. Zhu, *NNLO QCD Corrections to t-channel Single Top-Quark Production and Decay*, *Phys. Rev.* **D 94** (2016) 071501 [arXiv:1606.08463] [INSPIRE].

[153] ATLAS collaboration, *Measurement of differential cross-sections of a single top quark produced in association with a W boson at $\sqrt{s} = 13$ TeV with ATLAS*, *Eur. Phys. J.* **C 78** (2018) 186 [arXiv:1712.01602] [INSPIRE].

[154] CMS collaboration, *Evidence for associated production of a single top quark and W boson in pp collisions at $\sqrt{s} = 7$ TeV*, *Phys. Rev. Lett.* **110** (2013) 022003 [arXiv:1209.3489] [INSPIRE].

[155] S. Frixione, E. Laenen, P. Motylinski, B.R. Webber and C.D. White, *Single-top hadroproduction in association with a W boson*, *JHEP* **07** (2008) 029 [arXiv:0805.3067] [INSPIRE].

[156] F. Demartin, B. Maier, F. Maltoni, K. Mawatari and M. Zaro, *tWH associated production at the LHC*, *Eur. Phys. J.* **C 77** (2017) 34 [arXiv:1607.05862] [INSPIRE].

[157] CMS collaboration, *Observation of single top quark production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*, submitted to *Phys. Rev. Lett.* (2018) [arXiv:1812.05900] [INSPIRE].

[158] R.D. Ball and A. Deshpande, *The proton spin, semi-inclusive processes and measurements at a future Electron Ion Collider*, in *From My Vast Repertoire ...: Guido Altarelli's Legacy*, A. Levy, S. Forte and G. Ridolfi eds., pp. 205–226 (2019) [DOI:10.1142/9789813238053_0011] [arXiv:1801.04842] [INSPIRE].

[159] NNPDF collaboration, *Fitting Parton Distribution Data with Multiplicative Normalization Uncertainties*, *JHEP* **05** (2010) 075 [arXiv:0912.2276] [INSPIRE].

JHEP04(2019)100

[160] G. D'Agostini, *Bayesian reasoning in data analysis: A critical introduction*, World Scientific (2003) [INSPIRE].

[161] R. Boughezal, A. Guffanti, F. Petriello and M. Ubiali, *The impact of the LHC Z-boson transverse momentum data on PDF determinations*, *JHEP* **07** (2017) 130 [arXiv:1705.00343] [INSPIRE].

[162] J. Rojo, *Constraints on parton distributions and the strong coupling from LHC jet data*, *Int. J. Mod. Phys.* **A 30** (2015) 1546005 [arXiv:1410.7728] [INSPIRE].

[163] J.M. Campbell, J. Rojo, E. Slade and C. Williams, *Direct photon production and PDF fits reloaded*, *Eur. Phys. J.* **C 78** (2018) 470 [arXiv:1802.03021] [INSPIRE].

[164] D. Kraft, *A software package for sequential quadratic programming*, DFVLR, Köln (1988).

[165] M. Farina, C. Mondino, D. Pappadopulo and J.T. Ruderman, *New Physics from High Energy Tops*, *JHEP* **01** (2019) 231 [arXiv:1811.04084] [INSPIRE].

JHEP04(2019)100